Land use changes and potential impacts on environmental outcomes in Lam Dong province

A thesis
submitted in partial fulfilment
of the requirements for the Degree of
Doctor of Philosophy

At
Lincoln University
by
Lan Nghia Nguyen

Lincoln University

2016

ABSTRACT

Abstract of a thesis submitted in partial fulfilment of the
requirements for the Degree of Doctor of Philosophy

Land use changes and potential impacts on environmental outcomes in Lam Dong province

by

Lan Nghia Nguyen

Lam Dong is a province in the Central Highlands of Vietnam which has high biodiversity values. However, rapid socio-economic growth is causing environmental degradation. This research is directed at understanding the key driving factors behind individual choices about land use and the subsequent effects these may have on environmental degradation. Understanding individual motivations helps identify potential policy interventions and predict degradation with different policies. The research used a spatially explicit agent-based modelling approach (SeABM) to integrate spatial and non-spatial parameters in land use decisions. To determine the non-spatial parameters for the model, a questionnaire was used to collect demographic information from households. Spatial parameters were collected through data collection processes from different sources.

Based on the model that was developed, a number of simulations were done to evaluate the effect of different policy options compared to a base-line simulation. The effects were based on changes to four

key measures – land use changes, soil erosion, carbon dioxide sequestration and landscape fragmentation. The scenarios evaluated were reducing population growth rate, improving households' income, supporting low income households, promoting a perennial cashew crop, promoting acacia hybrid and promoting payment for forest environmental services. The baseline assumed that the population will grow at a rate of 2.7% per year as a normal rate. The number of mouths to feed in each household and the labours will also increase due to population growth. The cash balance of households may be deficit and lead to changes in their land uses to compensate for any loss. The land use changes will follow the historical trend described by several land use change modules results of analysing time-series satellite images.

Simulation outcomes with scenarios such as population growth, income growth and financial support did not deviate from the baseline scenario (the business as usual scenario). Moreover, they all potentially caused negative impact on the quality of the environment by increasing the amount of soil erosion, reducing the capacity to store more carbon dioxide in vegetative cover, and increasing the landscape fragmentation after 10 years.

Promoting perennial crops such as cashew in simulation had a positive effect on local livelihood. However, it still degraded environmental quality. Because this crop is usually planted at higher elevations with large spacing on steeper slopes, increased soil erosion results. Promoting acacia hybrid had a positive impact on livelihood and a lower negative influence on environmental outcomes. Acacia hybrid can be considered a good option as it is planted with higher density and has a higher growth rate compared to cashew, which helps reduce soil erosion on the ground and increase carbon sequestration. Providing payment for environmental services had a positive impact on environmental quality. The pressure on natural resources would be less if households had sustainable sources of income which encouraged them to protect and cultivate forest in the long term.

Based on the results of this research, several policy options would be suggested to harmonise both socio-economic growth and environmental quality. Land use conversion should be sustainably maximised if the current land use is not profitable for households. Long-term investment and long rotation for plantations at high elevation should be encouraged to reduce soil erosion and the restriction on rice land on flat areas needs to be flexible. Controlling the birth rate and improving education are also remedies to reduce the number of people depending heavily on agricultural activity and utilising forest products. Increasing incentive for forest protection and improving legal responsibility for violation in forest protection are highly recommended to encourage households to protect forest and sustainably benefit from it.

The results show that SeABM framework is capable of identifying land use dynamics in the future with different policy options. It has the potential to be used as a tool by policy makers in land use planning to explore new development alternatives. However, there are many gaps that can be improved upon with this research to have better modelling which captures the reality of land use decision-making behaviour.

Keywords: environmental degradation, forest degradation, land use planning, spatially explicit agent-based model, SeABM, GIS, Python, modelling, simulation, land use trajectories

TABLE OF CONTENT

ABBREVIATIONS

ABM	Agent-based Modelling
ADB	Asian Development Bank
CA	Cellular Automata
CAP	Center for Agricultural Policy Consulting (Vietnam)
CLUE	Conversion of Land Use and its Effects
DEM	Digital Elevation Model
EKC	Environmental Kuznets Curve
FAO	Food and Agriculture Organisation of the United Nations
FIPI	Forest Inventory and Planning Institute (Vietnam)
FRA	Forest Resources Assessment
GIS	Geographic Information System
GSO	General Statistics Office
IUCN	International Union for Conservation of Nature
LPSO	Lam Dong Provincial Statistics Office
LUDAS	Land-Use Dynamic Simulator
LULC	Land use/Land cover
MARD	Ministry of Agriculture and Rural Development (Vietnam)
MCDM	Multi Criteria Decision Making
NDVI	Normalised Difference Vegetation Index
NIAPP	National Institute of Agricultural Planning and Projection
OECD	Organisation for Economic Cooperation and Development
OLS	Ordinary Least Square
PFDP	Provincial Forest Protection Department
PFES	Payment for Forest Environmental Services
REDD	Reducing Emissions from Deforestation and Forest Degradation
SeABM	Spatially explicit Agent-based Modelling
SNV	Netherlands Development Organisation
SRTM	Shuttle Radar Topography Mission
UCLA	University of California, Los Angeles
UNESCO	The United Nations Educational, Scientific and Cultural Organisation
UN-REDD	The United Nations Programme on Reducing Emissions from Deforestation and Forest Degradation
VecGCA	Vector-based Geographic Cellular Automata

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

CHAPTER 1: ADDRESSING THE ISSUE

1.1. Introduction

Deforestation and forest degradation have become global issues that affect the livelihoods of millions of people living in developing countries. This awareness has been driven by remotely sensed data that have allowed for the tracking of deforestation over time. While deforestation defines the loss of forest areas, forest degradation refers to the destruction or reduction in the quality of specific aspects of forests.

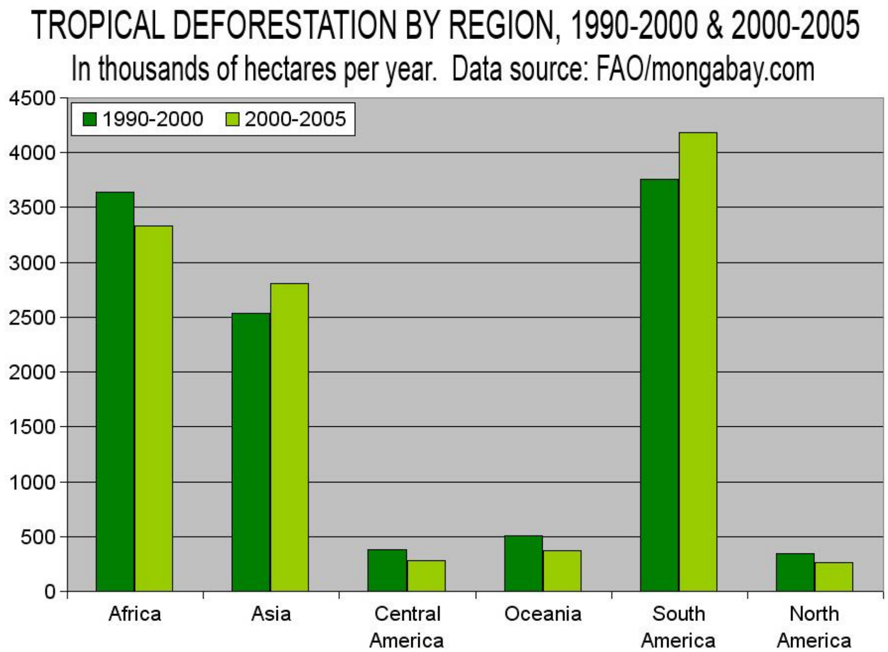
On a global scale, Butler (2006) suggested the most common cause of deforestation in the tropical region in the period of 2000 - 2005 was pasture land conversion. This activity was responsible for 40% of tropical deforestation. Subsistence agriculture and large-scale agriculture (soybean, oil palm, maize, and rice) caused 30% - 40% of the loss. Logging activities were estimated to contribute 5-8% of the total tropical deforestation while other activities like infrastructure, mining, urbanisation, non-agricultural fires etc. were the reason for 3% of lost area.

Permanent forest losses are linked to the corruption of government institutions (Koyunen and Yilmaz 2009) and the inequitable distribution of wealth and power (Schaeffer 2003, Koyunen and Yilmaz 2009, Mulvaney 2011). Rapid urbanisation due to high pressure from fast population growth in combination with overpopulation and uneven distribution has also been blamed for deforestation (Jepma 2014). As a consequence of those causes the global deforestation rate increased rapidly around the middle of nineteenth century (Wilson 2003).

According to the latest Global Forest Resources Assessment in 2015 forested area has decreased from 4,128 million ha in 1990 to 3,999 million ha in 2015, although the rate of the net forest loss has been cut by 50% (FAO 2015). The tropics, especially in Africa, Asia and South America, have had the biggest forest loss while most countries and territories in the temperate and boreal zones have shown a net increase in forest area.

Figure 1 summarises the deforestation in tropical region from 1990 through 2005.

Figure 1. Tropical deforestation by region, 1990-2000 and 2000-2005



As can be seen in Figure 1, Africa, Central America, Oceania and North America had downward trends in deforestation, where the area deforested in the 2000 - 2005 period is lower than in the 1990 - 2000 period. The average deforested areas in Asia and South America increased in the later period. By 2005, it was estimated that South America lost 4.2 million ha per year and Asia lost about 2.7 million ha per year (Butler 2006).

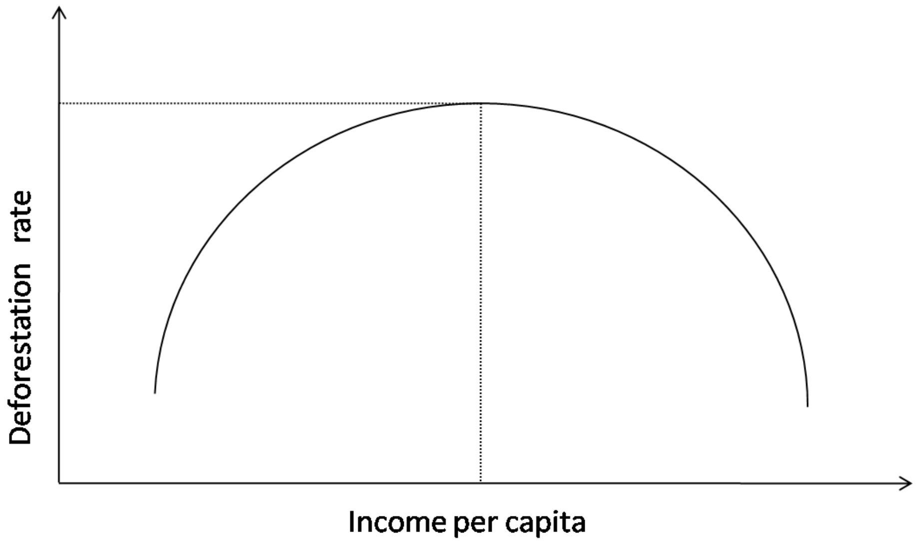
Vietnam is one of the Asian countries with high deforestation and forest degradation rates. As reported in the Global Forest Resources Assessment, in 2015 Vietnam had 14.7 million ha covered by forest (or 44.5% of Vietnam's total area) (FAO 2015). This is just a small recovery from a huge forest cover loss from wars and economic development between the 1960s and 1980s when the country lost two thirds of its forest cover, becoming the country with the highest deforestation rate in Southern East Asia. It was suggested that forest cover fell to as low as 20% of the country's total area in 1993 (De Koninck 1999, Newbery et al. 2000). Analysis of the available forest cover data by Meyfroidt and Lambin (2008a) shows that forest cover in 1991-1993 was 25-31% of total land area. It then then increased to around 32- 37% in 1999-2001. Asian Development Bank estimated that Vietnam lose about 185,000 ha of forest annually between 1976 - 1990 (ADB 2000). The reliability of figures reported by the Vietnamese government are uncertain as they may have inflated the values to argue that Vietnam had stopped illegal logging and deforestation, or they may also have been raised up to attract public attention on the deforestation rates (ADB 2000, Vajpeyi 2001, FAO 2003).

Despite being listed as one of ten countries with the largest annual net gain in forested area between 1990 and 2010 (OECD 2015), rich forest and medium forest areas in Vietnam declined by 10.2% and 13.4% respectively between 1999 and 2005 (Coe 2008). A report from the WorldBank (2005) indicated that in 2004 Vietnam had only 4.6% of natural forests classified as rich and closed-canopy forest and over two-thirds of the country's natural forests were defined as poor or regenerating forests.

Causes of deforestation in Vietnam are various. Forests were heavily affected during the Vietnam War (Vajpeyi 2001). However, since the war ended in 1975, the causes can be grouped into (1) high population growth, (2) agricultural extension, (3) timber harvesting for different economic sectors including construction and pulp industry (De Koninck 1999, ADB 2000). Swidden cultivation practices among ethnic minorities of Vietnam used to be blamed by Vietnamese government for the deforestation though both De Koninck (1999) and Lang (2001) did not agree with that assertion. Castella et al. (2005) found that in the 1980s, his study area lost 50% of its forest cover after a series of reforms in decollectivisation of agriculture in the northern uplands of Vietnam, where the land use rights of land previously owned and managed by the government was returned to individual farmers in the form of long-term lease (though ownership still belongs to the government).

Van Kooten and Wang (2003) examined the relationship between deforestation and income growth. Poor economies may take advantages of native forest exploitation to grow. However, increase in income may lead to demand in environmental amenities which urges those economies to reduce deforestation to improve environmental quality (Bhattarai and Hammig 2001). The trend of deforestation during an economic development phase could be explained by the environmental Kuznets curve (EKC) (Culas 2007). The EKC in this case was a hypothesized relationship between deforestation rate and income per capita, as shown in Figure 2.

Figure 2. Environmental Kuznets Curve



Source: Adapted from Culas, 2007

According to Figure 2, the growth of income per capital follows the increase in deforestation rate. Deforestation rate is highest when per capita income is in the middle range. Beyond this level of income per capita, the higher the income reaches the lower the deforestation rate is (Van Kooten and Wang 2003).

In many developing countries like Vietnam, deforestation seems to be the consequence of increasing income per capita. At some early stages of development, governments cannot avoid native forest exploitation and the gains it brings to economic growth thus increase the income per capita. The dependence on natural resources reduces when the economy reaches a certain level of growth. The EKC may be applied to describe the development of many economies like Vietnam. Good polices and management could reduce the height of the curve, and help nations reach a higher income per capita while slowing down deforestation rates.

Hai et al. (2013) suggests that the obvious consequences of long-term deforestation in Vietnam were increased erosion in upstream areas and siltation of irrigation systems in downstream areas. It also caused a series of serious floods and droughts in different regions of Vietnam (ADB 2000). Forests had been cleared for the development of large-scale hydropower projects and irrigation dams, People were forced to move from low areas to steeper areas to give place for reservoirs, and those people had to clear native forest on the slope for their livelihood, which later caused flood and erosion (Vajpeyi 2001). The destruction of mangrove forests along the coastal areas and estuaries of Mekong Delta provinces in the south of Vietnam in 1980s and 1990s for shrimp farming had weakened the natural coastal protection and led to loss of human lives and huge damage in typhoon Linda (Cosslett and Cosslett 2013).

Forests in some regions of Vietnam became more fragmented and subjected to edge effects due to the urbanisation and agricultural extension (Meyfroidt and Lambin 2008a), threatening biodiversity in Vietnam (Jamieson et al. 1998). The loss of forest in Vietnam also contributes to anthropogenic carbon dioxide emissions, which leads to global warming as suggested by van der Werf et al. (2009) for global scale deforestation and forest degradation. Vietnam has faced the challenge of reducing deforestation rates while also stimulating economic growth. When forest loss goes beyond the point of self-restoration, the economy may need to contribute more resources to remedy the damage. The desire to slow down deforestation and forest degradation is clear; however, it is constrained by unsustainable natural resource management. To pursue more sustainable development, Vietnam must address this conflict.

1.2. Research problem

Deforestation and forest degradation have become major issues worldwide, particularly for Vietnam. Considering the consequences of deforestation and forest degradation, the Vietnamese government has put its efforts into establishing and recovering forest cover since the 1990s. One of the early forest promotion programmes was the “Five Million Hectares” programme of reforestation launched in 1998, which intended to increase the 9 million ha of forest cover to more than 14 million ha by 2010 (MARD 2001). By increasing forest cover, the government expected to recover the environmental functions of forests in reducing erosion, flash floods and other unwanted consequences and gradually improving the livelihoods of forest dependant people. Along with reforestation, the reallocation of forestry land owned by the state government to households was carried out. This new tenure arrangement encouraged households to put more effort into improving forest quality with incentives and intensifying agricultural land rather than extending existing fields, which helped decrease agricultural cultivations on hillsides (Nagendra and Southworth 2009).

Vietnam has been actively participating in the incentive frameworks for developing countries such as the Reducing Emissions from Deforestation and Forest Degradation (REDD) programme, which emphasizes creating financial value for the carbon stored in forests, and the later version, REDD+ (or REDD plus), which was developed further to include the role of conservation, sustainable management of forests and enhancement of forest carbon stocks (UN-REDD 2015).

1.2.1. Research objectives

Despite the effort to slow down the rate of forest degradation by the state government, Vietnam still suffers the loss of its forest cover, especially in many highly forested provinces. In those provinces, households have been utilising the forest land for other purposes to meet their needs ignoring the restrictions on land use conversion. Forest lands are being converted into agriculture lands due to the demand for food in each household and this conversion has been driven by the market price for perennial cash crops such as coffee, rubber, cacao and pepper. Households usually act more quickly than the official planning directions from government for many reasons, which is why stopping deforestation and forest degradation seems to be challenging. The role of households in land use decision making is extremely important in the dynamics of forest landscape. Understanding the mechanism behind land use decisions could help policy makers to be more effectives in their efforts to introduce sustainable development. These facts urge the need to address the reasons behind the losses in quantity and quality of forests.

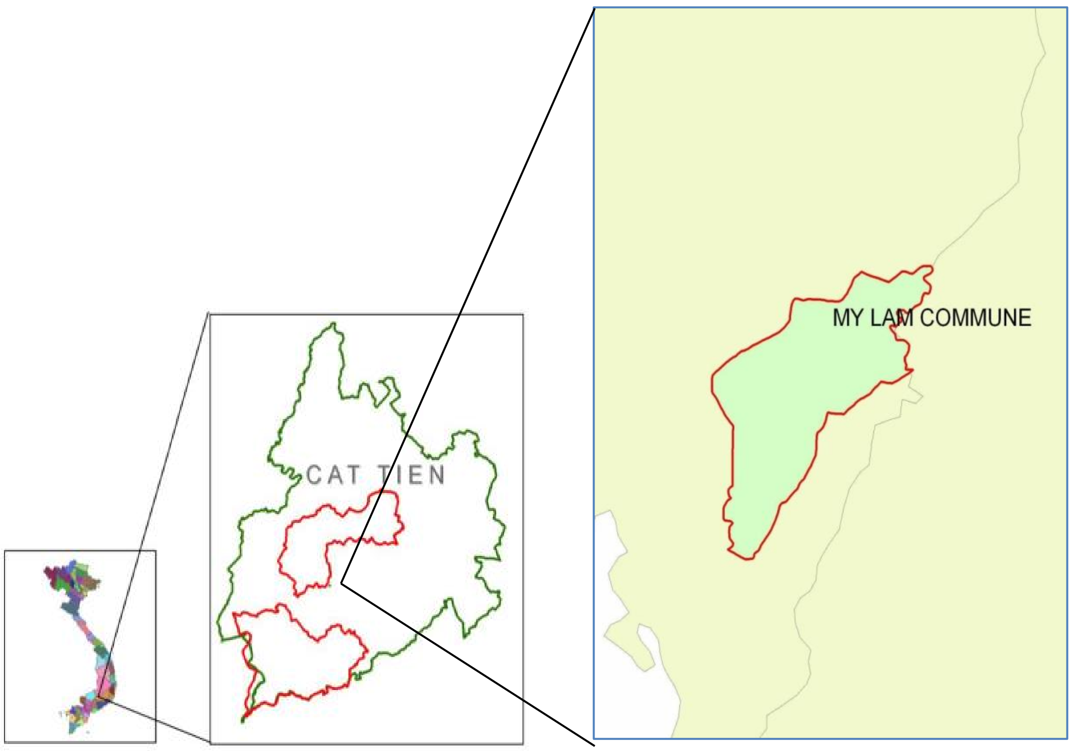
For this research it was hypothesized that environmental degradation is caused by a combination of biotic and abiotic factors which affect humans' interaction with the landscape and has a growing trend due to the rapid socioeconomic development of the Central Highlands region. Farmers make land use decisions based on a complex set of factors without foreseeing long-term consequences. Under the current hypothesis, this research aims to understand how farmers' decision making affects environmental outcomes which lead to environmental degradation, and how different policy interventions might influence that degradation.

Using My Lam commune in Lam Dong province as the study area, the main objectives of this research are (1) to identify the mechanism of decision making based on influencing factors; (2) to set up agent based modelling to simulate land use changes; and (3) to predict trends of future environmental outcomes, such as changes in forest cover, soil erosion and carbon sequestration. Moreover, the effects of different policy interventions such as alternation of income and population growth, or payment for forest environmental service, will be tested using simulation.

1.2.2. Study area

My Lam commune in Cat Tien district (Lam Dong province) has been selected as the study area given the scope of this research and budget constraints. Several criteria were considered when choosing the study area: proximity to intact or preserved forest, diversity of land use, pressure on land use conversion, and convenience to conduct the survey within budget constraints. Figure 3 shows the relative location of study area.

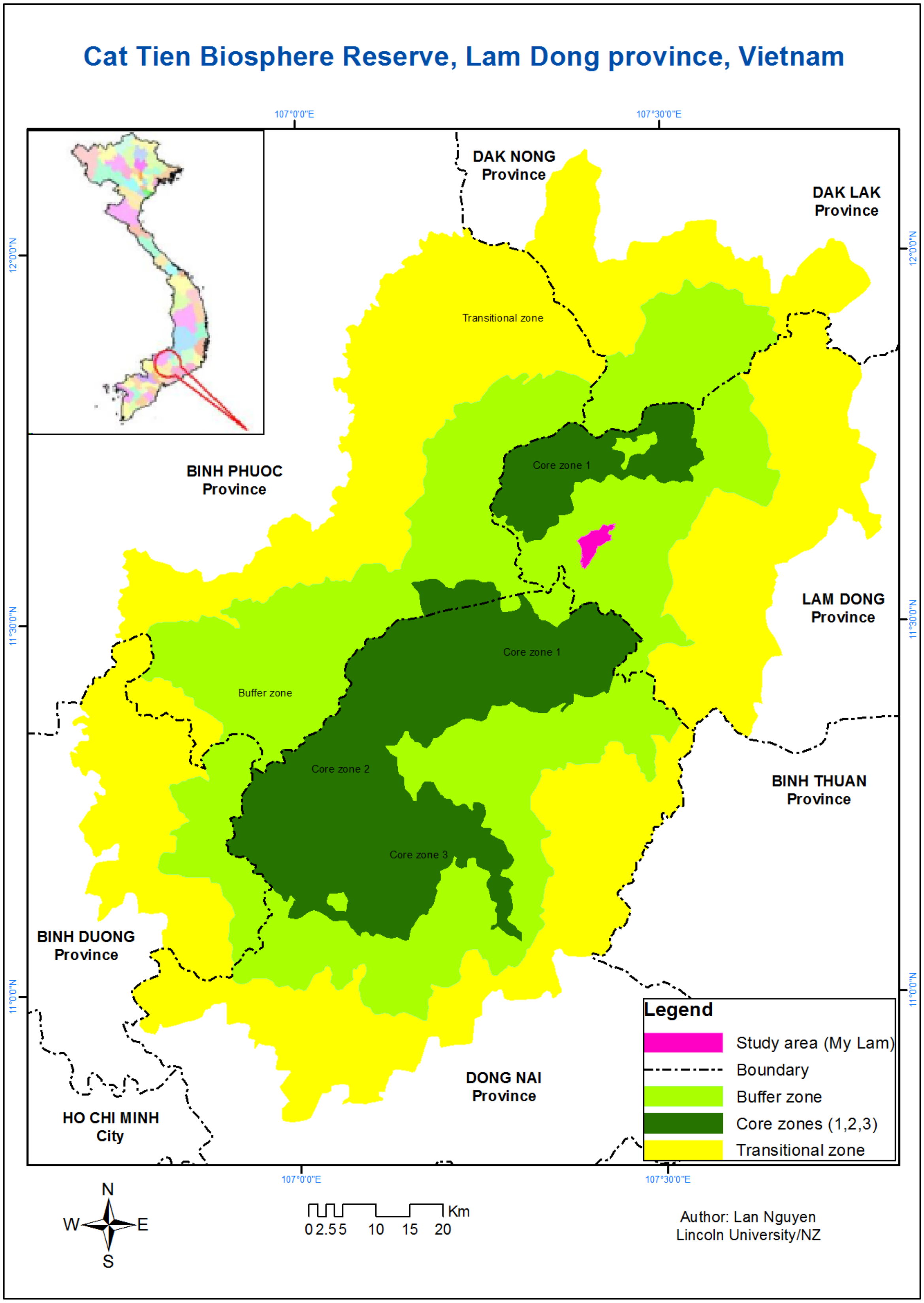
Figure 3. Relative location of My Lam commune in Cat Tien district and Lam Dong province



My Lam commune is located in proximity to the Cat Tien Biosphere Reserve with its population of approximately 1,181 people sharing an area of 1,568 ha in which 891.3 ha is forest land and the rest is agricultural land (Lamdong 2012).

The map in Figure 4 below shows that My Lam (filled with pink colour) is one of many communes situated in the buffer zone (in light-green pattern) of Cat Tien biosphere reserve and in the corridor which links two parts of the core zone (in dark green pattern). The buffer zone plays an important role in reducing negative external influences, especially from the transitional zone (area outside the buffer zone, in yellow pattern), on the quality of conservation in the core zone. However, due to the proximity to the core zone, any forest degradation and deforestation in the buffer zone may threaten the effort which has been made in the core zone, sooner or later.

Figure 4. Cat Tien Biosphere Reserve, Lam Dong province, Vietnam

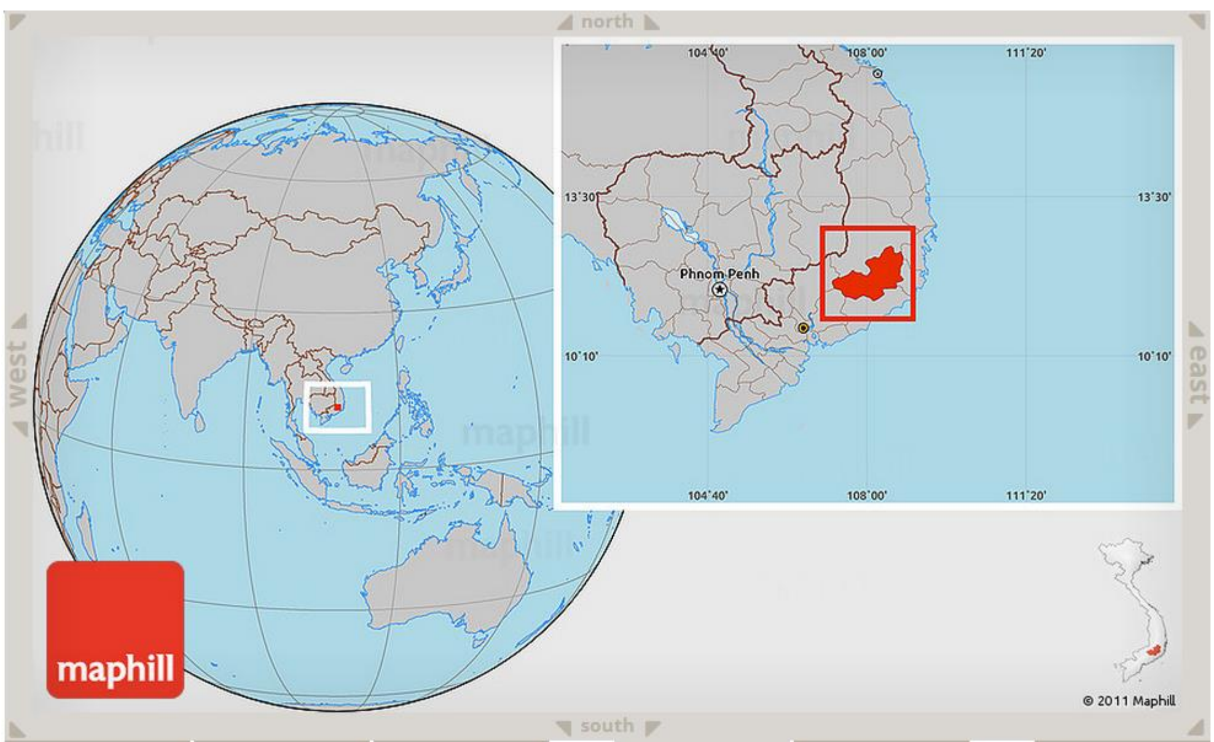


It was considered that land use dynamics, especially land use conversion to agricultural purposes, are very sensitive to the conservation activities in the core zone of the biosphere reserve (Huynh 2014). My Lam has average accessibility and can be reached by public transportation and motorbikes. Through a survey, demographic data and secondary data were obtained using interviews with a questionnaire. Collected data from My Lam commune were assumed to be representative enough to extrapolate the importance of land use changes and their consequence on environmental quality to a larger scale.

1.3. Lam Dong province and its environmental issues

The context of socioeconomic development and the land use dynamics in Lam Dong province have been summarised to provide an overview of background of the research problem. It helps to understand the rationale of why a specific area in Lam Dong province was chosen as the study area. Lam Dong is one of five provinces in the Central Highlands of Vietnam and is located in the southern part of Vietnam, as shown in Figure 5.

Figure 5. Map of Lam Dong province, Vietnam



Source: Maphill (2011)

Lam Dong province has a total area of 977,540 ha (BTNMT 2014), a forested area of about 60% (591,349 ha) a very high biodiversity value and rich environmental services (Lamdong 2010, UN-REDD 2013). It has a network of national parks and nature reserves with highly protected status to maintain the natural forest against illegal logging and to provide habitats for wildlife, flora and fauna. These forests are the habitat for many Red List species such as the Javan rhinoceros, golden-cheeked gibbon, bears, elephants and gaur (IUCN 2006). There are 99 globally endangered species and 110 protected species regulated by Decree 32/2006/NĐ-CP of the Vietnamese government found in Lam Dong's forests (UN-REDD 2013). Forest in this region strongly influences the Mekong river network, which has an important role in regional and trans-regional socio-economy, including hydro power development, agricultural irrigation, fisheries and water for urban use in nine downstream provinces (Lamdong 2010) These provinces include Ho Chi Minh City, the southern socioeconomic and political centre of Vietnam, and rapidly developing industrial zones like Binh Duong, Dong Nai and other provinces in the lower Mekong river basin.

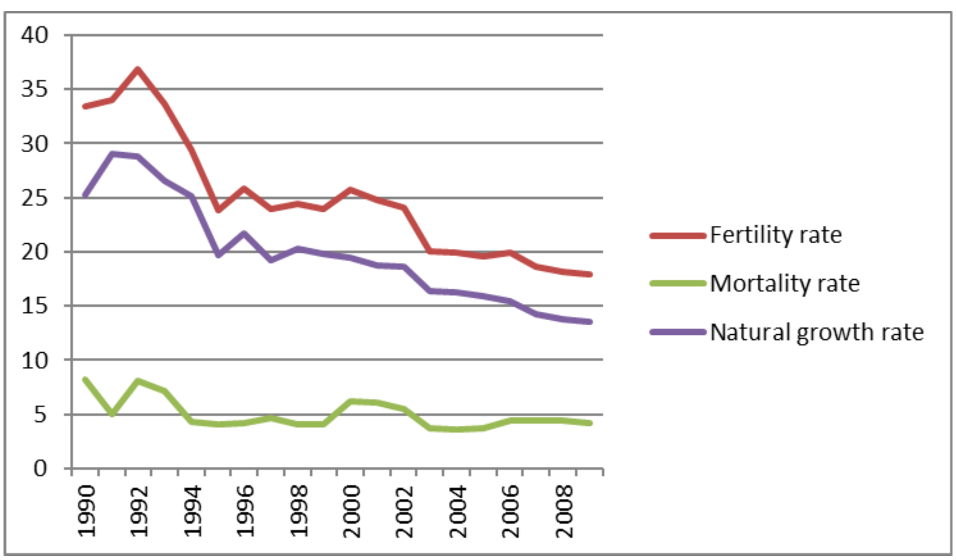
Rapid deforestation for economic recovery after the Vietnam War in 1975 and high immigration from low land Vietnam to the Central Highlands became a challenge for Lam Dong to maintain forest coverage due to illegal logging and forest conversion to other purposes. During 2008 - 2012 the Provincial Forest Protection Department (PFPD) had processed over 2000 violations mainly related to illegal logging, illegal exploitation of forest products, and deforestation for agricultural purposes (FPD 2013). In the first nine months of 2015, PFDP discovered 1,325 violations or 5 violations per day on average (Lamdong 2015). This shows an increasing trend and PFDP agreed that cases have become more severe and complicated. Thousands of hectares of natural forests in Lam Dong have been converted into cash crops such as coffee, rubber, pepper, flower and vegetables for export markets. The areas of loss in 2008 and 2009 were 2,569 ha and 1,926 ha respectively. About 25,000 ha of forests were also cleared for many hydropower plants in this province in 2005 (UN-REDD 2013).

In general, from 2000 to 2010, forest cover in the Central Highlands region reduced by 30% due to slash and burn and agriculture, burning in dry seasons, and illegal logging (Thiennhien.net 2010). Explanations for deforestation and forest degradation are varied but the main reason appears to be the fast growth of the population and unsustainable economic recovery in this region. Rapid population growth was mainly due to immigration (both planned and unplanned) and from the natural growth of immigrating populations (Zang et al. 2001). The planned migration was designed for both socioeconomic and political purposes while unplanned migration was blamed on shifting cultivation, family reunification or shortage of labour for timber and mining industries (Dang et al. 1997).

More than 80% of Vietnamese people are known as the Kinh ethnic group. They have a long tradition in paddy rice cultivation and populate the three main river deltas: the Red river delta in the northern part of Vietnam, the central coastal delta and the Mekong river delta in southern part of Vietnam (VN Embassy 2011). In order to reduce the pressure on cultivated land in the Red and Mekong river deltas and to promote mountainous rural socioeconomic status, part of the Kinh population has been resettled to the Central Highlands region, where other ethnic groups such as Ede, Rong, Sedang, Tai, and Giarai have been living in for centuries (Epprechtet al. 2007). These redistributions were popular in the 1970s - 1990s under the New Economic Zone Program (Zang et al. 2001). The net migration rate in the Central Highlands was the highest in Vietnam at 64.81% (Dang et al. 2003). During the 1990s there was an unplanned immigration of indigenous ethnic groups from the northern mountainous regions of Vietnam into Lam Dong province, primarily driven by high coffee prices (Epprecht et al. 2007). The average population density at the provincial level increased from 98 people/km^2 in 1998 (LPSO 2005) to 112 people/km^2 in 2009 with this immigration (LPSO 2009). At the district level, some districts like Don Duong and Duc Trong reached 154 and 184 people/km^{2}respectively in 2009 (LPSO 2009).

As a consequence of the rapid population influx, the region suffers a high rate of population growth, as shown in Figure 6.

Figure 6. Population growth rate in Lam Dong province



Source: Calculated based on LPSO 2005, LPSO 2006, LPSO 2007, LPSO 2008, LPSO 2009

As seen in Figure 6, recorded data show that the fertility rate in Lam Dong province had been higher than 20% for several years from the 1980s to 2000. Although this figure represents a downtrend of the natural growth rate in Lam Dong province, in 2008, it was 14% which was more than six times higher than the national average (Duong 2009).

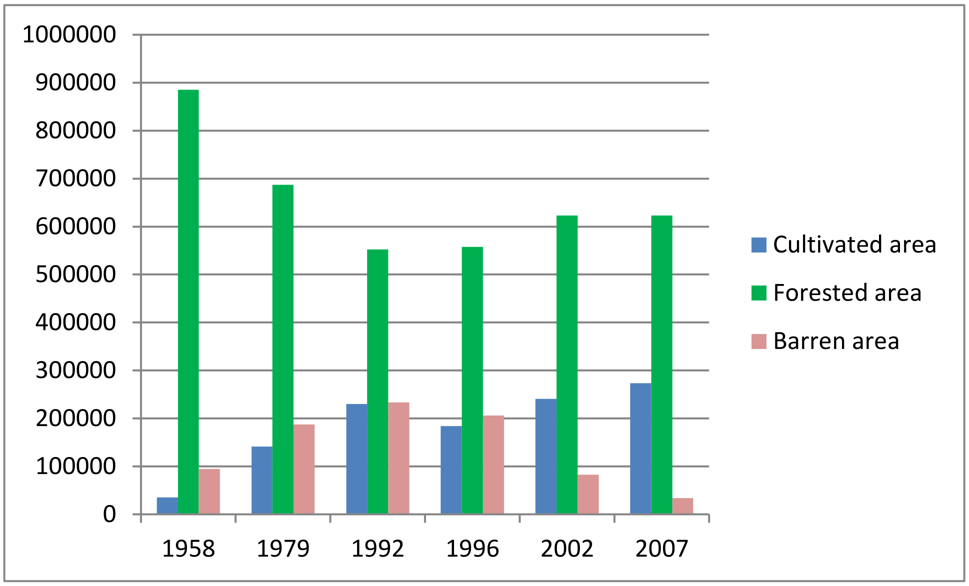
The local indigenous ethnic groups such as Ko Ho, Churu, Ede, Rong, Sedang, Tai and Giarai, who have traditionally practiced swidden or shifting cultivation and forest utilisation for centuries, found that forest resources could no longer sustainably support their lives under increasing population pressure (De Koninck 1999). The major change in ethnic composition, with a larger share of Kinh population, in Lam Dong has had several socioeconomic impacts. Other minority groups like the Hmong, Tay, Nung and Dao also migrated into this region and tripled the number of groups from 13 before 1975 to 40 groups at the present time (Hoan 2006).

A poverty and ethnic minority rate distribution study done by Epprecht et al. (2007) indicated that poverty rates and the distribution of ethnic minority groups in Lam Dong and the Central Highlands had a strong correlation. Indigenous groups often lost their traditional use-rights to their territories, as the government redefined and alienated property rights (Duy et al. 2010). Many groups gradually stopped their traditional livelihood practices to participate in the new commercial economy (Clarke 2001). Moreover, shifting cultivation practice among indigenous groups has had a negative effect on sustainability by increasing its frequency and intensity when cultivated land is limited. Forests have been felled for crops without consideration of natural regeneration and fertile soil preservation (Enright 2013). Forest resources in Lam Dong weren't heavily exploited until the 1970s. Research by Duy et al. (2010) using remote sensing analysis, showed a significant loss of forested areas between 1963 and 1984 in Lam Dong. Much reasonable old growth and high value forests had been chopped down during the economic recovery.

Forest degradation in Lam Dong province reflects a weak forest enforcement regime. De Koninck (1999) investigated agricultural land expansion in Lam Dong in 1958, 1979 and 1992 and found that the rate of expansion grew from 2% to 27% between 1958 and 1979. The share of cultivated land increased from 3.5% in 1958 to 13.9% in 1979, and then to 22.6% in 1992 in this province. About 60% of forest area cleared in this period was used for cultivation, 30% was left as barren lands and the remaining areas were in different forms of degraded forest (De Koninck 1999). Data analysis conducted by Sunderlin and Huynh (2005) revealed a negative correlation between high poverty rate and high forest coverage with low forest quality and poor forest management.

Data obtained by De Koninck (1999) using remote sensing techniques showed a decreasing trend in forested area from 1958 to 1992 while data published by the General Statistical Office of Vietnam (GSO) reports a recovering in forested area from 1996 to 2007. Figure 7 reflects a decreasing trend of forest coverage in the 1990s comparing to historical data in 1950s.

Figure 7. Dynamics of land use of Lam Dong province from 1958 to 2007 (ha)



Source: adapted from De Konnick, 1999 and GSO 1996, 2002 and 2007

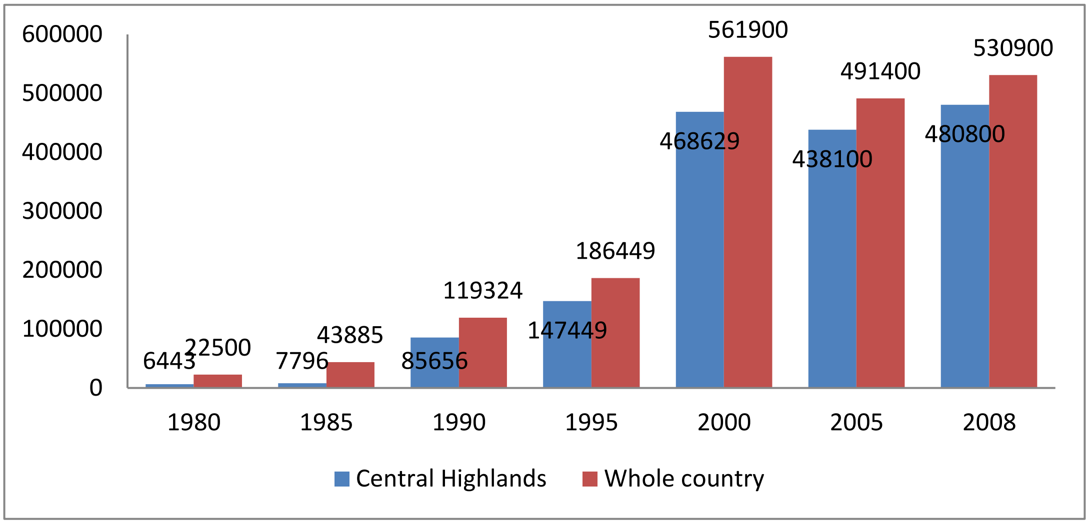
According to Figure 7, from nearly 900,000 ha of forested area in 1958, Lam Dong lost about 30% of that area in the 1990s. The forested area in Figure 7 increased from the 1990s to early 2000s and then held steady in the 2000s. However, it was about 200,000 ha less than what it had been in 1958.

In addition to population growth, road density and improved market access facilitate forest land conversion. If roads open up in formerly inaccessible areas, colonisation pressures can result in environmental degradation and forest conversion (Muller 2004). The establishment of the State Forest Enterprises with predominantly Kinh labours provided cheap forest timber and agricultural land for economic development, but it also put high pressure on the forest resource which later caused economic inefficiency and environmental disasters (Oscar 2003).

Agricultural expansion and intensification often coexist and occur simultaneously. While agriculture intensification has not reached levels that save land and labour, the cultivation expansion keeps continuing, with forest land being converted to perennial crops such as coffee, rubber and pepper. The promotion of perennial crops fully or partially sponsored by the government in Lam Dong induced massive land use change and widespread deforestation, particularly during the 1990s (Muller 2004).

According to the Ministry of Agricultural and Rural Development of Vietnam, a large area (including forest lands) has been being converted to perennial crops since the 1980s (SNV 2010). The Central Highlands region contributes a significant share to these conversions as shown in Figure 8.

Figure 8. Land devoted to coffee production in the Central Highlands (ha)



Source: Duy et al., 2010

As seen in Figure 8 the total area of land devoted to coffee plantation increased from nearly 6,500 ha in 1980 to approximately 480,000 ha in 2008. The income from coffee and other perennial crops encouraged people to extend cultivated areas by clearing forests rather than intensifying available planting areas. However, market fluctuations are always a challenge for farmers in this region. People hurry into investing and enlarging areas of perennial crops when prices go up and they clear to change to other crops when prices go down (Muller 2004, Duy et al. 2010, SNV 2010). The promotion of commercial crops by the government has brought more rapid land use and land cover changes by silently encouraging local farmers to quickly convert their suitable land into cash crops and to violate the conservation intentions.

Due to high prices, farmers converted their inefficient coffee and forest plantations to pepper, which could bring them a revenue of US$ 30,000 per hectare (Nam 2008). Even though this revenue is higher than other crops in this region, it is full of uncertainty caused by high establishment costs and unstable market demand. The population–land ratio increases the scarcity of cultivated land. The growing pressure on land encourages a shift to higher-value crops on existing fields and an expansion of agriculture into more fragile, marginal areas (Boserup 1965), which has high risk due to weak agricultural extension and irrigation.

The trend of forest degradation is likely to continue if under privileged people keep depending heavily on forest resources or they cannot find themselves better livelihood alternatives. Ethnic minorities who manage forests traditionally using indigenous knowledge have found difficulty in adapting to the economic development pace and in sharing the forest landscape with low-land migrants who prefer cash crops only. A recent report by UN-REDD (2013) estimated about 300,000 ethnic minority people (25% of Lam Dong's population) have been traditionally fully dependent on the forest. Different ways of forest exploitation regardless of forest regulations have exceeded the resource's recovery ability and resulted in poverty. Attempts to make land use more profitable in the short-term at a household level may affect the landscape in the long-term.

The consequences of initial forest cover losses driven by immigration and economic development bring about landslides and soil erosion, which results in changes to river geomorphology, desertification, release of carbon into the atmosphere and loss of habitats and biodiversity. Moreover, both agricultural expansion and intensification increase the use of agricultural herbicides, pesticides and fungicides and contributes to the degradation of air and water quality. A participatory research approach was employed by Ha (2001) to identify the environmental challenges in some typical communes in the Central Highlands where land use practices have had unwanted long-term environmental consequences on the landscape.

The annual rainfall in this region is high but distributed unevenly. About 70-80% of rainfall occurs in the rainy season (from early May to the end of October) and causes floods and land slips. The situation gets more serious when the forest coverage in upper hills and slopes is significantly reduced. Topsoil fertility is washed out heavily during the rainy season. In the dry season, the rainfall is only 15-20% of the whole annual rainfall thus resulting in serious drought.

Drought has been considered an environmental issue for the Central Highlands in research done by Tinh (2006). Lack of irrigation water reduced crop productivity and increased risk of investment loss. Drought is not only caused by natural phenomena but also by human activities which are seen as new factors significantly influencing global climate changes during the last few decades (Adger 1998). Moreover, the lack of adequate policies in many developing countries can cause drought or aggravate a drought situation (Hoc 2002). The expansion of crops boosts the demand for irrigation. Groundwater has been over-pumped for several years which exhausts water resources and increases the release of toxic contaminants (Muller 2004).

An irregular hydrological regime not only speeds up soil erosion but also spreads out agricultural toxic chemical residuals, which have been over used in this region. This directly affects the urban water supply on the downstream side. Moreover, water shortages in the dry season enable sea water to move inland, causing serious salinisation of river water and cultivated soil in coastal areas of Vietnam. Effects on soil conditions may be ambiguous depending on the applied techniques (use of cover crops or shelter trees) and the planted crops. The soil degradation in the Central Highlands was highly remarked in the Great Mekong sub regional report (Pehu 1998). Authors like Müller and Munroe (2005) and Ha (2001) considered the threats to environmental outcomes of this region as the trade-off between development and environment. Balancing economic development and environmental protection is currently one of the major challenges in the upland areas of Vietnam.

To slow down deforestation in the early 1990s the Vietnamese government initiated policies to encourage forest preservation. It had already issued several forest protection incentives to preserve forests, which included both political directions and financial support (or stipends). Currently, the maximum benefit for individuals from forest protection could reach US$ 8-9/person/month. The average income of poor household is US$ 55/household/month or US$ 12/person/month (Thuy et al. 2011) – just above the Vietnam national poverty line for the 2006-2010 period of US$ 11/person/month in rural areas (Prime Minister 2005).

Analysis conducted by SNV (2010) showed that the income of rural households came mainly from cash crops. This kind of income was increasing while incomes from forest resource utilisation, forest protection compensation and governmental allowances was decreasing. This situation encourages farmers to clear forest to expand their crops for higher income. Payment for Forest Environmental Services (PFES) is a financial tool being used for forested provinces in Vietnam, with the aim to promote sustainable livelihood for forest dependent people while functioning as forest protectors and forest environmental services providers. A PFES trial (Prime Minister 2008) carried out in Lam Dong was successful, with an average payment of VND 280 thousand/year/ha (approximately US$ 15) in reservoirs for Da Nhim and Dai Ninh hydropower plants.

It needs to be considered that, obviously, growth in population and associated necessities are sources of pressure on natural resources, and to satisfy the on-going demands agricultural areas were quickly expanded rather than deeply intensified and forced land use/land cover changes. The causes and processes of land use/land cover changes can be better understood by examining not only the physical patterns of change but also the factors influencing farmers' relationships with their farmlands (Boundeth et al. 2012).

Agricultural expansion with high immigration in the past has formed a large population dependent on agricultural activities as sources of income in the Central Highlands. The scarcity of available land for cultivation has become an issue for this province and due to that, forest land is likely to be sacrificed to fulfil the demand of land for agriculture and other purposes. The landscape has been being transformed continuously and so, land use and landscape transformations have had negative side effects such as environmental degradation and increased erosion. In any circumstances environmental degradation poses negative impacts and requires huge financial efforts and time to recover.

Land use decision making by local populations has been recognized as the main driving force in forest cover changes. There was a concern that land use decision behaviour is mainly determined by the economic situation and demography of households. It is always difficult to clearly quantify land use changes because they are complicated processes over time and space. However, they can be simplified as the result from a set of decisions made by land users (farmers, in this case). From this point of view, landscape transformation can be treated as the output of complex land use decision making made by farmers under the influence of many factors; by understanding the mechanisms of this decision-making process and its driving factors, government could more effectively address environmental degradation. This better understanding would help planners and policy makers to harmonise the forest conservation commitment with the improvement of local livelihoods.

1.4. Research questions

As described above, deforestation and forest degradation are believed to be the reasons behind environmental degradation in Lam Dong province. Forest losses are driven by several irrational land use practices and scarcity of agricultural land for high immigration. Land uses have been diversified and environment has been degraded across the landscape as a trade-off. The Kinh migrants highly preferred flat lands which were suitable for paddy rice and perennial crops, while other ethnic groups continued to cultivate steeper slopes. Land use fragmentation breaks the landscape configuration and results in a spatially heterogeneous land cover with consequences for quantity and quality of forest environmental services. To analyse the problem of how land use decision making influences the environmental outcomes, the following research questions will be addressed:

1. What are the main driving factors for environmental degradation in Lam Dong province?

2. What are the potential environmental outcomes without policy interventions?

3. What are the effects of policy interventions on environmental outcomes?

The first research question aims to find the potential driving factors which directly or indirectly affect environmental degradation through the decision-making process. This phase of the research investigates the nature and number of the driving factors, such as whether they are biotic or abiotic, and how strongly they correlate with environmental degradation. By answering the first research question, the research refines an understanding of the influence of the driving factors on environmental degradation and, more importantly, provides input information to a simulation framework, which is used to forecast the potential degradation in the future under different scenarios.

The second research question uses the findings from the first research question to build up the simulation framework to mimic how different driving factors work. This research question aims to model the simplest scenario where all the land use decision making by farmers happens as is, following a historic trend of decision making among farmers. The basic simulation framework serves as a baseline for further sensitivity analyses with other scenarios.

The third research question builds upon the basic simulation framework further by evaluating different policy scenarios. Those scenarios will be designed to simulate the effects of different policies which affect the livelihood of farmers. Answers for this question give several policy options.

Since the problem stems from the decision making behaviour for land use under the influence of different factors and its consequences on land use/land cover (LULC) change, an Agent-based model (ABM) with spatial integration is used to simulate the interaction among biophysical and socio-economic factors under (hypothesized) decision making behaviour of agents and mechanisms of their feedback to changes (Wainwright 2008). ABM allows for the dynamic representation of interactions between biophysical and socio-economic processes (Millington et al. 2008). To approach these research questions, this research needs to address the gaps in previous LULC researches and modelling in land use planning. The literature review in section 2.1 will present details of many studies sharing similar purposes with this research but implementing different methodologies.

CHAPTER 2: LITERATURE REVIEW

The research questions outlined in section 1.4 have oriented the literature review to focus on studies related to modelling and simulation of land use. The review starts with studies aiming to understand land use changes. Econometric modelling has often been used as a method for this aspect. This method was later empowered with the implementation of spatial data. However, to cope with the pace of land use changes it is necessary to have a robust approach to predict the changes accurately enough to assess their impacts on ecosystems and the socio economy of the territory of interest. The behaviour behind land use decision making is complicated due to multiple objectives with limited resources. At this point, the literature review focuses on Agent-based modelling research which incorporates individual behaviour in modelling land use dynamics. Land use decision makers are required to consider not only the scale of non-spatial factors such as soil types, cost, income, and benefit, etc., but also spatial aspects such as location and proximity, which can be captured by spatially explicit Agent-based modelling. By studying behaviour in a spatial context, the complexity of land use decision making processes can be better understood.

2.1. Review of land use land cover change modelling and simulation

2.1.1. Traditional approach

Land use change models based on economic models, trend analysis, and/or scenario analyses have been considered traditional approaches (Martine 2012). Econometric modelling of land use dynamics has a long tradition in using spatial information as independent variables as its estimation. An early model by Johann Heinrich von Thünen in 1826 represented the generalized land uses in the form of concentric rings surrounding a central market. Spatial factors were also combined with other non-spatial factors in land use decision making (Liu 2008). The models introduced by Chisholm (1962) and Cliff et al. (1997) became an initial but incomplete spatially dependent theoretical framework.

Recently, Boundeth et al. (2012) used an econometric model to estimate the influence of spatial and socio-economic factors on the probability of land use change from 2001 to 2007 in Northern Laos. The logistic regression model applied in this case has its advantages in terms of mathematical simplicity and meaningful results. However, the spatial factor in this research was a dummy binary variable which indicates if the distance from the land parcel of interest is within proximity of 500 metres to the road.

Martine (2012) explored the application of Multi-criteria analysis in modelling future land use scenarios for resource planning and management to support decision making in finding the best spatial allocation of land for future agriculture and forestry. This kind of Multi criteria decision making approach (MCDM) deployed a set of biophysical parameters known for influencing land use allocation such as current land use state, the Normalised difference vegetation index (NDVI), population, rainfall, elevation and road with their assigned weights in forming decisions (Martine 2012).

Traditional approaches seemed to focus on reasoning the causes of land use changes while modelling accompanied with simulation methods developed by the 1990s had tried to disentangle the cause-effect relationship between land use changes and their related driving forces at different scales. Land use change modelling under dynamic and complex interactions of factors and drivers has been simulated using the Conversion of Land Use and its Effects model (CLUE-s) (Verburg et al. 2002). This model uses empirical analysis to identify the relationships among spatial land use distributions, other driving forces and constraints of land uses, then determines the competitive advantages of those land uses and spatially simulates their allocations (Verburg et al. 2002).

The development of GIS and remote sensing has brought about more availability of spatial data for land use modelling. There has been interest for many researchers such as Yongjin et al. (2010) and Takenori et al. (2010), to analyse land use patterns using time-series satellite images for predicting future land use patterns using the Markov Chain based probability model with construction of spatio-temporal transition matrices. Transition matrices have been used in landscape ecology and GIS studies of land use to quantitatively estimate the rate of change based on a set of images and processed in GIS environment.

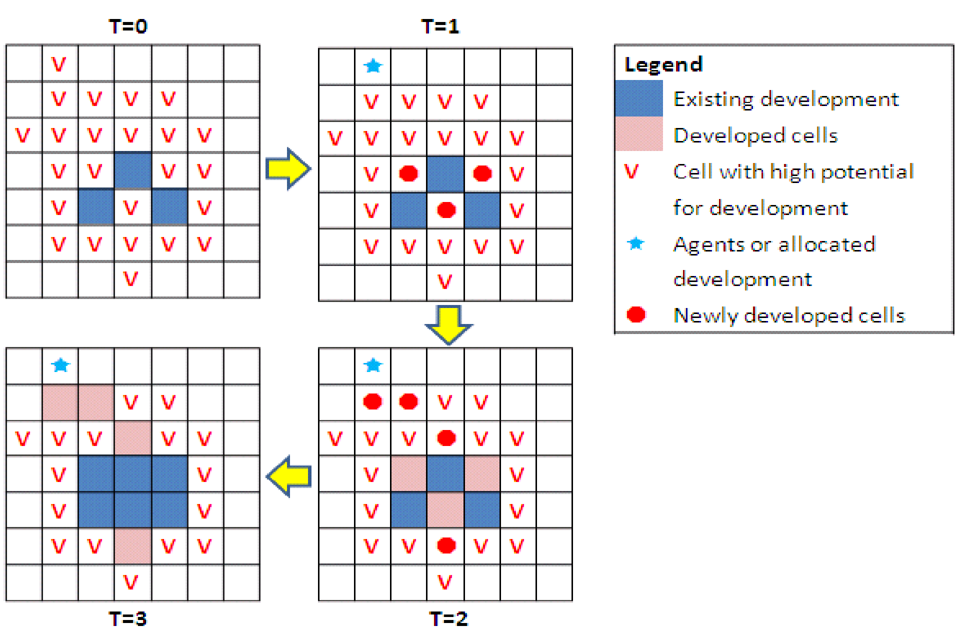
The role of spatial data has been increasing from just simple forms of independent variables in econometric modelling to complex interactions among different land uses in CLUE-s modelling. However, traditional approaches did not consider human behaviour when making land use decisions because this is a difficult metric to be included in a linear manner. The influencing power of spatial factors may be underestimated without considering the behaviour of the entity who makes decisions. Incorporating individual behaviours in modelling land use changes is necessary to make the land use decision more natural. For this purpose, Agent-based modelling seems to be an appropriate option.

2.1.2. Agent-based modelling and simulation

The complexity of human geography has been exhibited in different systems from land use change and population segregation to urban sprawl as well as intercity interactions and it is difficult for many types of modelling to capture this complexity. Both cellular automata (CA) and Agent-based models (ABM) are classes of computational models for simulating individuals' or collective entities' actions, behaviour and interactions and they offer some of the simplest frameworks to approach complex systems (Clarke 2014). The details of these two models are represented below.

Cellular automata (CA) is an early typology used to connect agents to each other. CA is a discrete model which is one kind of geographic simulation method (Yu et al. 2009, Filatova, Verburg et al. 2013). The basic framework of CA is a regular grid of cells; each cell has a finite state at time t. The state of a cell will be updated using predefined rules and its neighbourhood cells' status through time step t+1. This is an implemented form of spatial ABM which uses a "grid" environment to represent the agent's environment. Each cell is an agent and has a state which depends on the state of its neighbours. Conway's Game of Life (Gardner 1970) is good example of CA. Figure 9 presents an example of how CA simulation progresses with an area of two-dimension grid bounded in 7x7 cells or a regular space tessellation. CA which uses this kind of space tessellation is called raster-based CA.

Figure 9. Simulation of cellular automata modelling



Source: adapted from Candau et al., 2000

The graphical representation in Figure 9 shows that at the initial state (T=0) (upper left) of the grid some cells were identified as "Existing development" (or existing urbanised areas) and "Cell with high potential" to become urbanised areas. After some time, at state T=1 (upper right), an urbanised cell was allocated on the grid in a random non-urbanised cell (the blue star). The presence and the location of the new urbanised cell influenced the state of the grid through the rule set. At this point, some potential cells stay close to "Existing development" or existing urbanised areas change their status to "Newly developed" urbanised areas, thanks to their proximity to existing urbanised areas. At the later moment, T=2 (lower right), those "Newly developed" cells or new urbanised areas update their status to "Developed" urbanised areas while, under the pressure of population growth, other high potential cells close to urbanised zones become "Newly developed" urbanised areas with a simple rule: getting as close to the urbanised areas as possible. At the final state, time T=3, continuous growth of residential areas through the simulation process has linked the new urbanised areas with existing development blocks. The further the evolution is, the higher the number of cells being turned to existing urbanised area (Candau et al. 2000). In Figure 9, cells change their status in each step toward a "development" state during the simulation process. Moreover, a spatially-explicit model like this is used to simulate the changing processes at a fine scale in order to support the investigation of spatial patterns of humans' activities on a landscape scale, by exploring interactions between a coupled human –environment system over time (Deadman et al. 2004).

Agent-based models (ABM) are able to take into account an individual's (agent's) behaviour in land use decision making (Matthews et al. 2007). ABM and simulation was introduced in the 1990s to model complex systems of interaction and autonomous agents and has been being developed rapidly not only in the scientific domain but also in the commercial domain (Tobias and Hofmann 2004). ABM focuses on modelling the heterogeneity of agents and Sugarscape (Epstein and Axtell 1996) is a good example of an ABM. ABM can be based on theories of an agent's behaviour for contexts or situations in which the agents are involved. There are a number of theories and empirical heuristics related to agent behaviour in optimising, utilising, learning or making decisions (Macal and North 2010).

A typical ABM has three elements as described in Macal and North (2010):

- A set of agents and attached attributes and behaviours

- A set of relationships and methods of interaction which define how agents connect or interact with each other and with their environment

-The environment in which the agents interact

An ABM progresses by having agents repeatedly execute their behaviours and interactions over a timeline which has a time-stepped, activity-based, or discrete event simulation structure. In land use change research, these process-based models link the feedback from different decision making agents with a cellular model of the physical landscape to represent the influence of an agent's decision making behaviour on land use patterns, with attempts to interpret in a spatially-explicit manner across a range of scales (Parker et al. 2003). The understanding of agents' decisions in ABM models can be defined by bridging the gap between context and generalisability (Janssen and Ostrom 2006).

In comparison to traditional econometric modelling approaches, ABM has many differences. Firstly, in ABM, heterogeneous agents are modelled individually and instead of trying to maximise their welfare or optimise their outcomes as in the traditional approach reviewed above, ABM agents consider only their local environment to achieve possible and satisfied utility (Macal and North 2010). Secondly, each agent in ABM has its own individual behaviour (from average behaviour) rather than assumed rational behaviour (Castle and Crooks 2006). Thirdly, the involvement of agents' interactions with heterogeneous behaviours make ABM more complicated than the traditional approach which is full of simplistic assumptions with uncertainties (Hamill and Gilbert 2015).

The applications of ABM vary widely from business and marketing (Arthur et al. 1997, Macal 2004a, Macal and North 2010) to epidemic prediction (Bagni et al. 2002, Carley et al. 2006) and military applications (Hill et al. 2006, Moffat et al. 2006). The results of modelling and simulation are often used to inform policies and decision making. Some ABMs focusing on the conversion of intact tropical forest to other non-forest land uses can be found in the researches of Evans et al. (2001), Deadman et al. (2004) and Huigen (2004). The conversion of forest land to other covers plays an important role in accessing the environmental outcomes in areas of interest. In these researches, ABM models try to extrapolate the local management rules to a whole landscape using input data such as land uses in the past, accessibility to land, soil characteristics, and some local statistical variables such as population, ethnicity and livestock. Simulations mimic the agents' (farmers') decisions on allocating resources such as labour, land, capital, activities and cultivation. The designed scenarios help stakeholders to assess the potential impact on natural resources.

Since the potential applications of ABM can be quite diverse, there has been increasing interest in ABM and developing its methodologies. Authors like North and Macal (2007), and Marsh and Hill (2008) have discussed the design and implementation environments for ABM. A bottom-up approach with highly interactive design methodologies seems to be effective in practice (Macal and North 2010).

CA and ABM share some similarities. They are both a bottom-up approach which is built from elemental actions or interactions at microscale level which then is aggregated over time and space to form behaviour at macro level, and these models come with simulations to dynamically evolve the system to mimic the real world. The core difference between CA and ABM is that the latter is considered to have the ability to deal with more behaviour units which interact with each other over time and space (with the integration/coupled with a GIS) and can perform objects' movement but not the limited space defined by grid as in CA, to represent mostly the change of cells' value.

The latest modifications of raster-based CA in combination with ABM offers an open structure with modifiable inputs or adjustable parameters (Castella et al. 2005, Chen et al. 2009). This approach allows users to alternate the modelling inputs with various planning functions or scenarios to adapt to different scenarios. It increases the scope of application of the GIS simulation and makes it practical for decision-making. Soares-Filho et al. (2002) developed a complicated simulation called DINAMICA which uses a stochastic cellular automata model in its core. This is an application of logistic regression to calculate the spatial dynamic transition probabilities and simulate landscape dynamics in the Amazon. However, raster-based CA models have been seen to have cell size and neighbourhood configuration sensitivity affecting the outcomes of simulation (Ménard and Macal 2005, Kocabas and Dragicevic 2006). In other words, changing the cell size of a grid may affect the results and speed of simulation. Larger cell size may have different simulation outputs from a smaller cell size. Many authors like Couclelis (1985), Takeyama and Couclelis (1997), White et al. (1997) and O'Sullivan (2001) worked on reducing the scale sensitivity of raster-based CA by recommending to choose the right cell size to fit research objectives. The recently proposed vector-based geographic cellular automata model (VecGCA) has overcome the scale sensitivity issue by allowing an irregular space tessellation. This is also a two-dimensional surface made from almost any kind of geometric shapes rather than regulars shapes. With VecGCA the neighbourhood definition and the transition functions are associated to the real properties of each geographic object within the study area, and that allows the geometric transformation of the geographic objects as a result of the transition functions (Moreno et al. 2008). The definition of the 'influence function' is an important part in the modelling approach. As described by Moreno et al. (2008), an influence function can be defined in Equation 1 as follows:

Equation 1:

$$\boldsymbol{g_{ab}} = \begin{cases} \boldsymbol{1 - e^{-\propto_{ab}} if \ 0 \leq \propto \leq \propto'} \\ \boldsymbol{e^{-(\propto_{ab} - \propto_{ab'})} if \ \propto > \propto'} \end{cases} \boldsymbol{0 \leq gmax \leq 1}$$

g_{ab} is the influence of the neighbour “a” on the object “b”, if g_{ab} is higher than a threshold then state of “b” changes toward state of “a”

$\propto_{ab}$ is a determinant factor of influence of the neighbour “a” on the object “b”

$\propto_{ab'}$ is the value of $\propto_{ab}$ on the border

At each time step, the model will check the influence intensity on every object and its value will be compared to a threshold value; based on a predefined set of transition rules, a part(s) or whole object may be transformed to another state. The output of each time step will be the land use map for (t+1) in vector format. However, the disadvantages of this approach are long processing time and the dependence on rasterization in geometric transformation. Those are gaps for that could be improved in future.

As reviewed above, land use decisions by humans is a complex process influenced by natural, socio-economic factors. However, both traditional modelling approaches and cellular automata have not explicitly modelled the human-environment interaction, especially how humans consider spatial and non-spatial factors in their land use decisions. A spatially explicit Agent-based land use model (SeABM) can be a flexible way to describe the potential land use changes determined by human behaviour in space and time with consideration of socio-economic and spatial factors.

The combination of GIS and a spatial model such as CA is a typical application of a spatially explicit ABM. GIS topology provides a rich set of important geographical information related to agents. This combination with raster layers in GIS can be applied as a decision making tool in planning tasks and for predicting dynamic land use changes, as suggested by Clarke et al. (1997) and Engelen et al. (1999). Application of this combination was also found in urban sprawl research (Wu 2002), socio-spatial dynamics (O'Sullivan and Torrens 2000, Samat 2007), forest fire spread and soil desertification (Yu et al. 2009).

ABM and GIS can have a loose-coupled or tight-coupled integration. While the first type needs to import/export data between stand-alone ABM and GIS software, the second type has the ability to process modelling data and GIS operation within one application or minimise the import/export between applications. GIS solutions such as ArcGIS are classified as tight-coupled integration thanks to the built-in spatial simulation model builders like macro language, Avenue language, Python scripts or Geo-algebra (Takeyama and Couclelis 1997). Loose-coupled integration is a more flexible approach which tries to fulfil the insufficiency of tight-coupled integration in representing complex land use simulations (Clarke and Gaydos 1998). This integration uses other programming languages (C++, Java, etc.) and exports to GIS packages to represent and analyse data. Examples can be found in research by O'Sullivan and Torrens (2000) and Castella et al. (2005).

Irwin and Geoghegan (2001) were interested in explaining the causal relationships between individual choices and land use change outcomes while reviewing some of the advances that have been made in developing econometric models of spatial land use. Brown et al. (2005) discussed deeply how ABM and GIS can be coupled. They pointed out four key relationships which affect how the geographic data and the ABM interact: identity, causal, temporal and topological. The identity relationship indicates how an agent in ABM associates with its geographic data; it serves a key link between attribute data to geographic representation. The causal relationship allows an agent to affect the spatial object and its value even if the spatial object is not associated with the agent. How an agent updates its attributes over time is controlled by the temporal relationship. Attributes can be altered synchronously or asynchronously depending on the configuration of the models. The topological relationship manages the consistency and the movement of spatial objects in GIS.

Applications of SeABMs should start with the determination of potential variables or metrics and their measurements. Later, the mechanism of land use decision making will be worked out through measuring and weighting factors affecting it. The decision making behaviour will be translated into a "protocol" to be used as the core of simulation (Parker et al. 2003).

The common questions for any land use change models including SeABM could be: (i) When do they make land use decisions? (ii) Where will they make the land use decisions? (iii) How will they make the land use decisions? Based on these questions several variables or metrics for modelling can be addressed. For a spatially explicit ABM, McGarigal and Marks (1993) proposed that landscape metrics such as patch distribution, number of patches, patch size and mean perimeter-area ratio explore and quantify the spatial patterns of a landscape structure. A less complex and less patchy land cover fragmentation could improve the ecological functions of the forests and bring better economic prospects, in terms of economic scale when fields are less fragmented or dispersed. Agents' decision behaviour is strongly related to non-spatial information, such as availability of cash and labour at the household level (Rounsevell et al. 2010). Those factors will influence the decision by an agent to change a land use to a desired land use state, if the rule set deems it is more suitable. Millington et al. (2008) suggested biophysical metrics (initial land use status, production capacity, location, etc.); economic metrics (profit, market demand, financial sources, etc.) and social metrics (ages, agent's world view, neighbourhood, etc.). Other socioeconomic metrics should be considered when exploring an agent's behaviour, including sources of income (farming or non-farming, subsistence or not subsistence, dependence or not on resource exploitation, etc.), road accessibility (based on Euclidean distance to main road types), wealth indicators, accessibility to credit, schooling, etc).

According to Templeton and Scherr (1997), Kaimowitz and Angelsen (1998), and Pender et al. (2001), agents in an ABM should react to the dynamics of different metrics and adjust their decisions according to the new circumstances. For convenience, authors like Brondizio et al. (2002), McCracken et al. (2002), Castella and Quang (2002) and Deadman (2004) classified decisions on land use as occupying (buying/being allocated a plot), releasing (selling/abandoning a plot), and maintaining/converting land use purposes. This simplification helped the modelling process

For a SeABM in land use change, it is important to develop a set of rules in the land use conversion process. This will guide which land uses can be converted in what circumstances, regarding not only the benefit of land use but also the biophysical conditions and especially the land use regulations of current local authority and country. This kind of hierarchy was discussed in Deadman et al. (2004). At the top of this hierarchy, the land use decisions for agriculture and forestry made by a household were driven by questions like “Are subsistence requirements for the household met?” or “Does the household have enough labour to deforest?” A poor household is willing to deforest for cultivation if it has enough labour to do so. However, it must leave the land as fallow if it does not have enough capital and labour for annual crops or if the soil is too poor to grow. McCracken et al. (1999), Brondizio et al. (2002) and Deadman et al. (2004) successfully defined land use trajectories based on household composition in its stages. Those authors found that nuclear-young (with small children) and nuclear-adult (with older children) households have higher potential for deforestation to set up livelihoods while in later stages they may become involved in plantation and agroforestry.

Along with different approaches in land use modelling, simulation has been developed to inherit the results from modelling to predict LULC changes over different future time periods. Simulations with the ability to run scenarios are useful for sensitivity analyses where different inputs are evaluated based on their corresponding results from simulation framework. Simulation tools such as Swarm RePast (Collier 2000) have made dynamic simulation technologies accessible to researchers in a broad range of disciplines, including social sciences (Deadman et al. 2004), computing science, economics and ecology (Minar et al. 1996).

In Vietnam, the application of ABM related to forest land use dynamics is not common and still limited to classical raster-based CA. Research in participatory simulation of land use change with the combination of ABM with an integrated role-playing game and GIS (SAMBA-GIS) was conducted in Northern mountainous region of Vietnam to determine the reasonable allocation of resources such as land, labour force and capital (Castella et al. 2005). In this research, the authors have deployed several tools such as a narrative conceptual model, an agent-based spatial computational model, a role-playing game and a multi-scale GIS to understand the interactions between humans and natural systems. Those combined tools synthesized (i) farmers' strategies (individual decision-making processes as a function of the farm's resource profile), (ii) institutions that define resources' access and usage, (iii) dynamics of biophysical and socioeconomic metrics. The simulation process used an annual time step with the support of GIS to represent the changes to the landscape due to decision making by farmers (agents). The land use simulation was based on the principle that agents allocate their resources (land, labour, and capital) to different activities (cash crops, husbandry, forestry and non-farming activities) which suit them best.

Loi et al. (2010) approached the land use suitability analysis in Lam Dong province. This was a couple between CA and GIS using various criteria, including natural capacity of a land unit and the implications of the socioeconomic, environmental and geographical features, in assessing the suitability of land use.

Castella et al. (2005) stressed the importance of using role-playing games to eliminate the effect of "institution" on agents' decision behaviour. Agents were asked to manage their resources as they wanted, not as planned or following what they were told to. They also indicated that an ABM approach should start with a narrative story of a local area to better understand the decision-making behaviour, and the model should be gradually refined through interactions between researchers and local people to get the desired result. In this research it was found that ethnicity and cultural aspects had little effect on the land use decision making.

Muller and Zeller (2002) investigated the effect of geo-physical, agro-ecological, and socio-economic determinants on land use change in Dak Lak province (one of five provinces in the Central Highlands of Vietnam) and assessed the influence of rural development policies on land cover change. The multinomial logit model was used to estimate the influence of hypothesised determinants such as socio-economic, policy, rainfall, soil suitability and topography on land use and the probabilities that a certain pixel has one of five land classes classified by satellite images from the years 1975, 1992 and 2000. The collected data were spatially referenced using geographic information systems (GIS). In this study, the behaviour of all land users and other stakeholders in land allocation and management had been aggregated at the village level (Muller and Zeller 2002).

Le (2005) carried out empirical verifications of the Land-Use Dynamic Simulator (LUDAS) model's components and its application to Hong Ha watershed in Central Vietnam. LUDAS is a multi-agent system model for simulating spatio-temporal dynamics of coupled human–landscape systems to assess policy impacts on landscape and community dynamics. In this model, decision-making mechanisms of households regarding land use were represented by spatial multi-nominal logistic functions and heuristic rule-based techniques. This sophisticated approach used typological livelihood to define land-use decision behavioural patterns of households who interact in a system of landscape environment, characterising individual land patches with multiple attributes. A set of policy factors were determined for land-use choices by households. Households' characteristics, environmental and policy information were integrated into land-use decision making procedure by agents (Le and Park 2008).

So far, the review of land-use change modelling studies has revealed that the application of ABM has become popular. Spatial data have been used in traditional econometric approaches and they have been more useful in ABM. The explicitness of spatial data is highly considered in land-use modelling. Households' behavioural decision making is the centre of ABM, however, its integration into a modelling framework is a complicated process. To balance between fulfilling objectives and constraints of this research, it is necessary to develop a flexible application of ABM which is spatially explicit and easy to deploy.

Clarke et al. (2014) stressed the important parameters to assess an ABM. They are design, performance, traceability, validity and calibration. The design or structure parameter needs a clear definition of scale, type of agent, action and behaviour used in ABM. The computing power and algorithm are measured by the performance parameter. The traceability parameter tries to make sure the ABM matches the actual or predicted outcomes. And if the prediction made by ABM is accurate or acceptable then it meets the validity parameter. The calibration parameter helps finetune the model to meet the performance, traceability and validity based on its initial design.

Land uses are not spatially uniform; they vary according to the position and configuration of land use types and other factors. For example, not all forest areas and forest categories in Lam Dong produce the watershed protection function, only forests with some specific criteria. Both spatial and demographic factors may affect individual behaviour in making land use decisions. This research needs a Spatially explicit ABM (SeABM) approach which allows incorporation of spatial and non-spatial data in one modelling framework and offers a flexible way to explore and simulate the environmental outcomes of different land use decisions across the landscape. Advantages offered by SeABM cannot be found in other reviewed approaches.

2.2. Land use trajectories

Land use and land cover change (LULC) has been named as a major cause of environmental change, not only on a local but also on a global scale. This information is based on the fact that nearly one-quarter of the tropical rainforest biome has been somehow fragmented or removed by human needs (Wade et al. 2003). It is hard to identify the mechanism behind land use cover changes, but the patterns of those changes can be identified with time-series data.

The consequences of deforestation and forest land conversion were discussed in many aspects, from the global carbon cycle (Cramer et al. 2004, Levy et al. 2004) and climate change (Laurance 2004, Mayle et al. 2004) to biodiversity (Sala et al. 2000, Gaston and Spicer 2004), health (Patz et al. 2004) and culture and economy (Godoy et al. 2005). They could be studied through the investigation of land use trajectories which resulted from a long-term and complex interaction among populations and their land use decision making under the influence of different causal drivers and factors.

LULC trajectories could be described as the temporal sequence of LULC classes at the pixel level, obtained through classified satellite images assembled in a time-series (Mena 2008). LULC trajectories can be used as tools to assess the links between policies, development programs and management implementations to the consequences of land use. Mena (2008) showed that there are different disciplinary approaches to obtain different ways of quantifying LULC and sketching up the trajectories. These procedures can be classified into four groups based on their methodological complexity:

(a) simple measures of change, such as the proportion of change: this is a widely used method showing the simple difference of areas or proportions of a certain class or classes, sometimes divided by time to obtain the rates of change within a certain area and land class between two points in time without being spatially-explicit;

(b) annual rate of change: this method involves the calculation of the annual rate of change assuming linear or non-linear changes between end-member periods. The Food and Agriculture Organization (FAO) calculates the annual deforestation rate using compound interest formulas to capture the annual exponential discount within the forest class between two moments in time, but also without being spatially-explicit;

(c) changes in the spatial arrangement or structure of LULC: this approach is based on landscape ecology and quantifies the changes in spatial configuration and composition of LULC at multiple dates and between dates of change;

(d) spatially-explicit representations of change at the pixel level using the concept of pixel histories: in this approach the spatially-explicit representations of change generated when two maps or raster objects are compared (using pair-wise comparisons) to obtain a change between two overlaid points in time, recorded as a single surface;

This research used the latter approach to characterize the LULC trajectories, which focuses on the change of land use types, their spatial configurations and the probability of change over time and space using the classified satellite images taken from 1995, 2000, 2005 and 2010. According to Gustafson (1998), the spatial configuration of land use change is based on the paradigms of landscape ecology, supported by the strong interconnection between the spatial pattern of landscape and its ecological functions and processes. Based on the results from the analysis of LULC trajectories, general land use changes will be set up to serve as land use change modules in the modelling framework. Land use change modules are the pre-set types of land use change from one kind of land use to another based on its history of change. This content can be coded in simulation scripts to activate the land use change when it is needed.

An interdisciplinary methodology combining the utilisation of processed satellite images and local expertise consultation was used. There are two steps in this methodology:

(1) identifying and analysing the LULC trajectories at the pixel level for the study area with the intention of building up the land use change modules

(2) allocating the land use change modules to the profiles of household decision making

The first step will be processed as in section 4.3 and the second step will be analysed in section 4.4.

CHAPTER 3: METHODOLOGY

The simulation developed in this research covers a range of sectors from social knowledge to land use management practice and requires an interdisciplinary approach. There are different methods and techniques deployed to answer the research questions outlined in previous chapters. They range from simple descriptive analyses to complex econometric modelling and simulation. There are also different data formats required for those methods and techniques which can be grouped into two sub-groups: spatial data (including geographical information and georeferenced data) and non-spatial data (including socioeconomic data). All the methods, techniques and data work together within a common SeABM simulation.

3.1. Choice of econometric model

For the first research question, the econometric model plays an important role. It helps define the driving factors of environmental degradation and how they influence degradation. The econometric model is needed to evaluate how different demographic factors (or non-spatial factors) and geographical related factors (or spatial factors) affect land use cover changes in the study area. The human induced land use cover changes can be used as a proxy of environmental degradation because forest cover losses from converting forests into agricultural land or residential areas may cause carbon loss, soil erosion or landscape fragmentation. This econometric model is assumed to be representative enough to explain the reasons behind land use changes in the whole province.

The spatially-explicit data were used as estimators (or independent variables) in the econometric model to reflect the influence of geographic parameters on the model's output. Every household has its own characteristics which are taken into consideration while making decision. Those characteristics were also estimators of the econometric model.

In general, households make land use decisions under the influence of different factors, and this routine is the main causal reason for land use changes. The econometric model focused on answering the questions: What are the causes of the changes? And how strong are they in the overall influence on land use change? In other words, the mechanism of decision making was driven by a combination of landscape features and household demography.

This econometric model was the tool used to quantify the correlation of spatial and demographic factors with land use cover changes in the study area in decisions made by households. By interpreting the output of the econometric model, the driving factors of environmental degradation could be identified. Coefficients of independent variables specified the conditions needed for any land use changes.

It was a challenge to choose a regression model and adequate variables to explain the diverse and complicated land use dynamics in this landscape. Thanks to the availability of multi temporal land use maps from 1995 to 2010 (derived from satellite images) there was a simple way to quantify the change of a predetermined period by using binary surfaces. A binary surface was obtained by overlaying a pair of land use cover maps of the same area at different years and comparing pixel's land use values at different years. If the land use values at a pixel are different from year to year, then the difference is coded as "1"; alternatively, if there is no difference it is coded as "0". By applying this rule to all pixels in the pair of land use maps it will generate a map of "0" and "1" values or binary surface. Logistic regression was identified as an appropriate approach to deal with binary values (Long and Freese 2005).

Stata software was used for model estimation and offered a family of logistic regression methods to choose from. In this research, probit regression was selected because it fits the requirement to be able to deal with the regression where the dependent variable can only be one of two values: "0" and "1".

Probit regression is a non-linear estimation method and it has been one of options to deal with dichotomous dependent variables (Park 2009). In this type of regression, the inverse standard normal distribution of the probability is modelled as a linear combination of the predictors. The regression produces a list of the log likelihoods at each iteration for a probit model. This kind of regression uses maximum likelihood estimation, which is an iterative procedure. The first iteration (called Iteration 0) is the log likelihood of the "null" or "empty" model; that is, a model with no predictors. At the next iteration (called Iteration 1), the specified predictors or dependent variables are included in the model. At each iteration, the log likelihood increases because the goal is to maximize the log likelihood. When the difference between successive iterations is very small, the model is said to have "converged" and the iterating stops.

Since the observed dependent variable, Y, is binary, the linear estimation approach cannot be used to estimate the coefficient of predictors (or independent variables). In this case the probit model with binary response directly describes the probabilities $P(y_i = 1)$ of the dependent variable y_i . In the Short Guide to Econometrics, Schmidheiny and Basel (2014) explained comprehensively how to transform a binary response model into linear model. N is a set of observations that is independently and identically distributed and has the dependent dummy variable y_i ($i = 1,...,N$) and a (K+1)-dimensional vector x_i' of explanatory variables including a constant. If (K+1) = 3 means the model has 2 explanatory variables and a constant, the probability that the dependent variable takes value 1 is modelled as:

$$P(y_i = 1|x_i) = F(z_i) \ = F(x_i'\beta)$$

where β is a (K + 1)-dimensional column vector of parameters and $z_i = x_i'\beta$ is a single linear index.

The single linear index then is mapped into [0,1] space using the transformation function, F, and satisfies the following general conditions:

$$F(-\infty) = 0, F(\infty) = 1, \partial F/\partial z \ > \ 0$$

The probit model assumes that the transformation function F is the cumulative density function (***cdf***) of the standard normal distribution. The response probabilities are then:

$$P(y_i = 1|x_i) = \Phi(x_i'\beta) \ = \int_{-\infty}^{x_i'\beta} \emptyset(t)dt \ = \int_{-\infty}^{x_i'\beta} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \, dt$$

where $\Phi(x_i^{'}\beta)$ is the probability distribution function (***pdf***) and $\emptyset(t)$ is the cumulative distribution function (***cdf***) of standard normal distribution. The cumulative distribution function indicates the probability that the variable takes a value less than or equal to x_i.

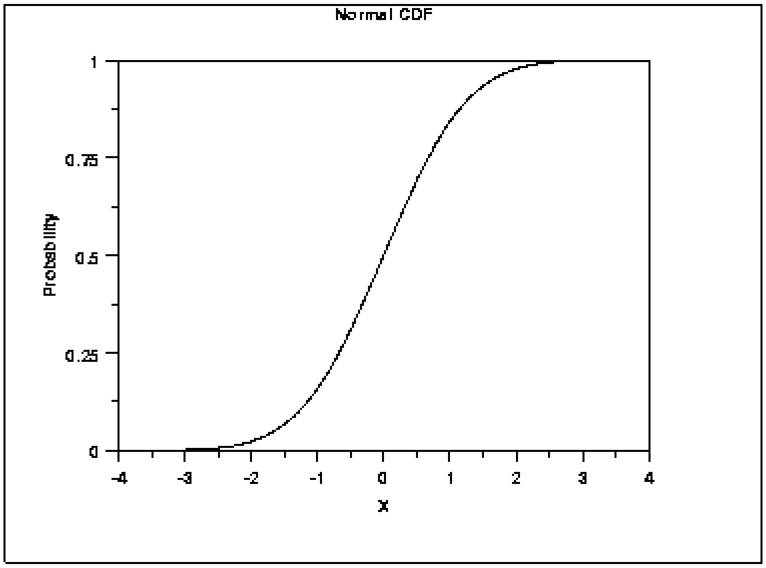
At this point, the probit model used to predict the dependent variable, Y, can be converted to linear as in Equation 2 below.

Equation 2:

Y* = *F*(Y)

The meaning of this kind of transformation can be explained as in Figure 10 below.

Figure 10. Cumulative distribution function of a normal distribution



Source: NIST/SEMATECH 2015

The horizontal axis (or x axis) is the allowable domain for the given probability function. Since the vertical axis (or Y* axis) is a probability, it must fall between zero and one. It increases from zero to one as we go from left to right on the horizontal axis. This transformation using ***cdf*** is applied in section 5.1.2 to calculate the values of Y*, which will be used to evaluate the probability of land use change.

The output of the econometric model, or the probit regression in this case, can be considered as the core of the SeABM which then simulated land use cover change throughout space and time. As discussed in section 3.1.3, any PolyCell needs a value representing its probability of being changed (or Y* value), which is specified by landscape features at location of interest and by households' characteristics and this value indicates the readiness to be changed to other land use of the PolyCell. This value then will be compared to a permanent value (Z value) which characterises the resistance to change the land use type at each PolyCell of interest in the landscape. It is assumed that a decision of land use made by a household could be realised if the readiness to change value is equal or higher than the resistance value (or Y* value is equal or higher than Z value).

3.2. Spatially explicit Agent-based modelling

3.2.1. Conceptual model

A conceptual ABM for the current research was adapted from the work of Parker et al. (2002), Deadman et al. (2004), Moreno et al. (2008) and Rounsevell et al. (2010). The simple form of the model is stated below.

land use change = Agent decision (spatial, non-spatial)

In this equation an "agent" represents a household and it shows that the **land use change** at a location, managed by that household, is the output of a decision-making function made by agents with parameters that are the set of spatial and non-spatial factors associated with that location. Since the agent's decision in the proposed model is spatially explicit, the model can be considered as a SeABM.

The challenging point of this research is how to define the agent's decision-making function, which has spatial and non-spatial factors as inputs. As a requirement of this research was to identify the dynamics of LULC, the potential approach should have the SeABM framework with a decision-making function based on the econometric model in its core. The simulation part of this framework should be carried out on a yearly time step for a certain time span and the agents in the simulation have the ability to drive land use changes after considering a set of spatial and non-spatial factors in their decision making. The approach should also be able to test different scenarios to compare the outcomes of different changes in policies. The outputs of this approach should be ready as inputs for other models such as carbon sequestration, erosion and landscape fragmentation.

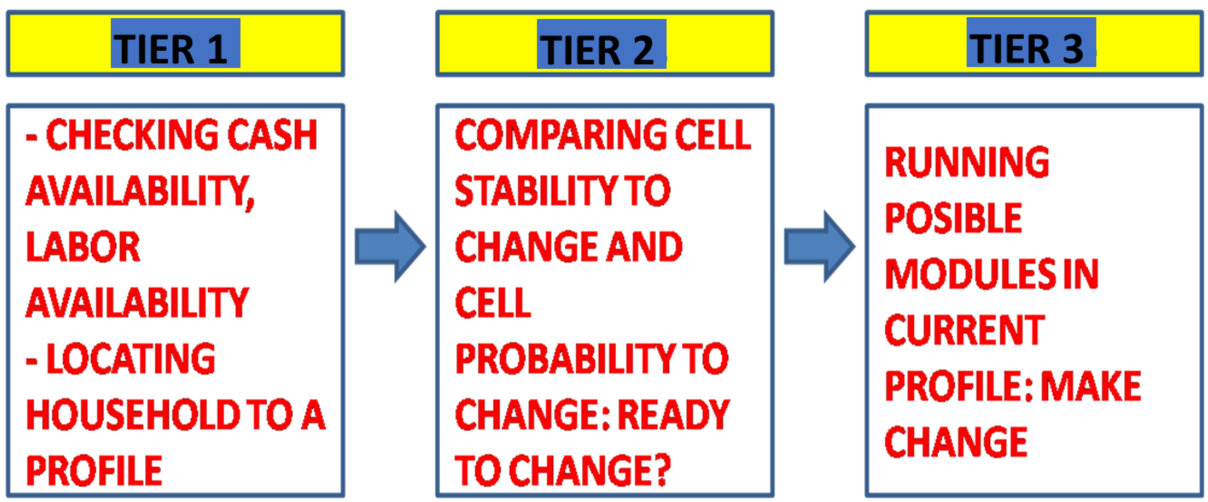
In comparison to the reviewed models in section 2.1, this SeABM simulation will be tightly coupled with GIS to take advantages of geoprocessing functions and spatial analysing power and perform the landscape transformation at the end of the simulation in the format of land use layer.

Due to the constraints of time, budget and manpower, My Lam commune has been chosen as the study area. This study area was assumed to be applicable to the research questions and could be used to draw an upscaled picture about land use dynamics and environmental degradation in Lam Dong province. How this study area was chosen, and its characteristics are well defined in section 1.2.2.

3.2.2. Hierarchy of decision making

Decision-making behaviour is a complex process in which an agent must consider different factors at different stages. In this research, the decision-making process has been fitted into three tiers of a hierarchy. On one hand, this hierarchy was designed to simplify the real decision-making flow and on the other hand it works coherently with the SeABM approach in this research. The order of this hierarchy is from the left to the right as represented in Figure 11 below.

Figure 11. Hierarchy of decision making at household level



In Figure 11, at the first tier (Tier 1) of decision hierarchy, households were grouped into different household profiles due to their cash and labour availability. There are four household profiles based on the balance of cash and labour availability in households. Households with surplus in cash balance and labour were allocated in **Profile 1,** where households have more than enough resource to change the land use, invest in expensive crops like cashew, or convert poor forest to more profitable land use. When the cash balance was in surplus and labour balance was in deficit, the household had cash but no labour availability to get involved in labour intensive activities. This is **Profile 2,** where households kept low labour demand activities and could invest in higher profit crops. In **Profile 3,** households did not have cash but they had spare labour. In this situation, households could carry out some low cost but labour intensive activities, such as converting poor or bamboo forest to other land uses to get quick return. In the last profile (**Profile 4,**) households were lacking both cash and labour. They had limitations for investment into high cost crops such as perennial crops. These crops could be abandoned or converted to annual crops by households if they were able to afford a low investment and low intensity labour. Using household profiles helped narrow down the decision-making options of different households with an assumption that cash and labour availability were the first factors households would consider before forming any idea of how a LULC decision could be made. These profiles characterise how an agent (or a household) could alternate LULC by using different LULC change modules.

At Tier 2, the consideration of different factors prior to a LULC change was converted into a comparison between the stability of any LULC changes and the probability of being changed at different points of interest on the landscape surface. The mathematical model was assumed to realize the complexity of human behaviour in evaluating different factors against each other. This kind of comparison was used as a proxy to show the willingness to change a LULC by an agent.

When a point on the landscape surface is ready to be changed, then at Tier 3 an agent has to follow the LULC change modules in its given profile. In other words, after evaluating all factors and finding that a LULC could be changed, an agent would change the LULC type in accordance to its demographic characteristics.

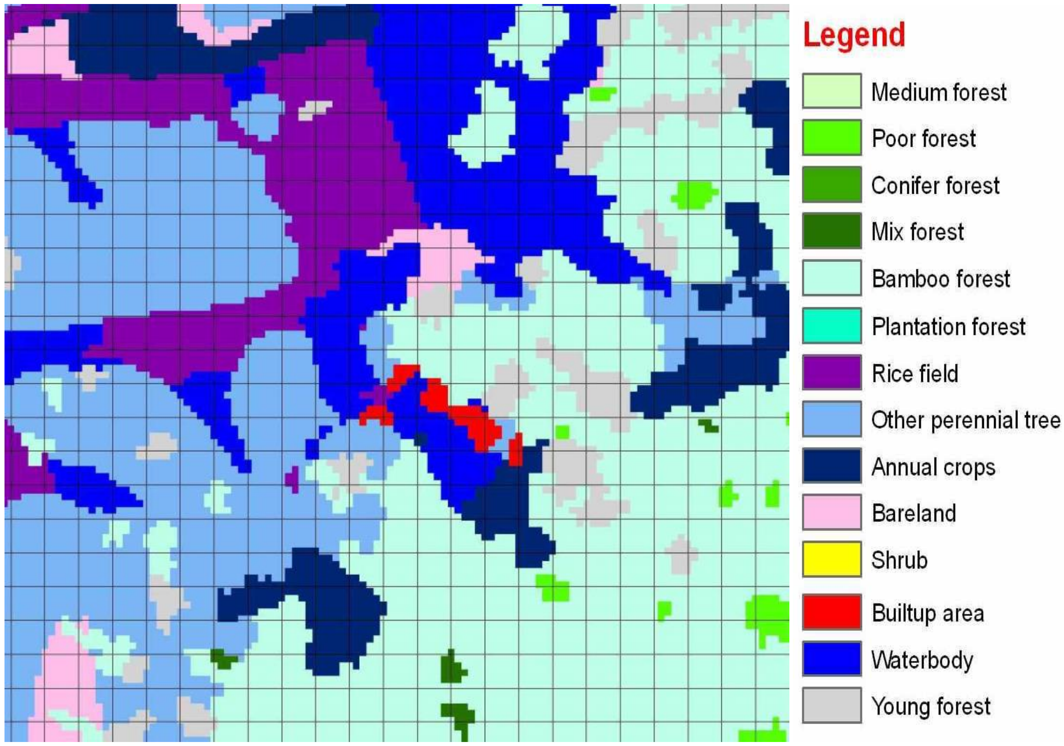
3.2.3. Design of PolyCell

Processing the spatial data during simulation is critical. To avoid the drawbacks of raster-based CA and VecGCA, the simulation core needs to handle the spatial data more flexibly to not only reflect the transformation of landscape when LULC change but also to reduce the processing time. The idea of constructing a PolyCell (Polygon Cell) was introduced in this research to satisfy those requirements. Each PolyCell is a square polygon.

After digitising the land use allocation sketch map into a feature class, a new vector layer was created using the Fishnet command in ArcGIS 10.1 (ESRI 2013) to produce a net of 20m x 20m square polygons. The choice of the PolyCell size follows the common spatial resolution of satellite images used in this research because some spatial features will be extracted from images to PolyCells' attributes. The procedure of this extraction is represented in the section 3.2.3.4

Each PolyCell can carry many attributes because they are vector objects. Those attributes can be inherited from the land use allocation map or extracted from different rasters. At this point, the PolyCell is considered as the basic spatial unit of the spatially explicit ABM. The representation of PolyCells is displayed in Figure 12 on top of a layer of land cover.

Figure 12. Representation of PolyCells



In Figure 12 a PolyCell may cover a whole land use type of the land use map underneath. In this case its land use attribute will carry that type. For other cases where a PolyCell may cover more than one land use type, its attribute will take the type under its centroid. A household (or an agent in SeABM) can "own" many PolyCells and this depends on the area allocated for the household in the land use allocation map, and attributes of one or several PolyCells can be changed under the decision-making process. Figure 13 represents a snapshot of attribute data of a PolyCell.

Figure 13. The data structure of PolyCells

Attributes of ML_extr_final

	CellID1	No_mem	Farm_inc	MLbuf_net_ctr	MLbuf_net_	Balance	MLbuf_net_	MLbuf_net_	MLbuf_
	17637	7	502	1584	1303	783	550	0	
	16933	7	502	1584	1303	783	800	0	
	16859	7	502	1584	1303	783	180	50	
	16244	7	502	1584	1303	783	424	150	
	14974	7	502	1584	1303	783	112	0	
	14738	7	502	1584	1303	783	650	550	
	14137	7	502	1584	1303	783	750	350	
	13854	7	502	1584	1303	783	50	250	
	13690	7	502	1584	1303	783	255	391	
▸	13237	7	502	1584	1303	783	381	50	
	13241	7	502	1584	1303	783	472	50	
	13098	7	502	1584	1303	783	335	532	

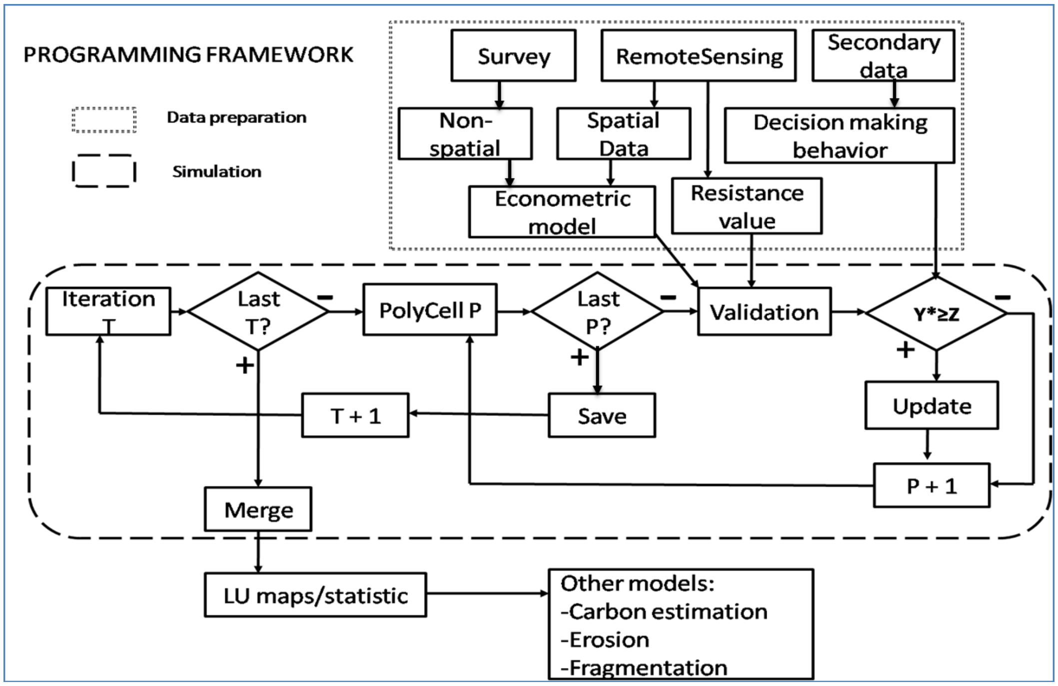
This is a typical attribute table of a polygon feature class in ArcGIS. Columns are fields which hold data from a survey using a questionnaire and many data processing steps which are detailed in section 3.2. In the GIS environment, processing the PolyCells may not be as fast as raster cells, however, PolyCells offer more advantages than raster cells thanks to holding many attributes favourable for modelling purposes and performing the geometric transformation.

For this research, it is also advantageous to manipulate the PolyCells in the modelling framework using Python scripts coupled with ArcGIS. The construction of PolyCells has an important role because based on this basic unit, several types of analyses, modelling and simulation were carried out. This vector datatype is easy to manipulate in GIS environment.

3.2.4. Programming framework

The programming framework represented in Figure 14 below explains the general workflow used to construct a SeABM for this research. According to this framework, the whole data manipulation, processing and analysis was divided into two blocks. In the first block, called Data Preparation, all necessary data were extracted, manipulated and analysed to serve other further steps in the Simulation block.

Figure 14. The programming framework for simulation



As seen in Figure 14, transition rules were constructed based on the spatial and non-spatial information that an agent had to balance out when making decisions on land use. Agents mentioned in this modelling process were households living in the study area who have direct or indirect involvement into forestry and/or agriculture. Agents' behaviour is supposed to affect the landscape dynamics.

The Data Preparation block prepares and defines important inputs for the Simulation block. Demographic data from a field survey and data from remote sensing materials were reclassified as spatial and non-spatial factors and were then fed into an econometric model as explanatory variables which help to measure the influences of different driving factors on land use changes. i.e., agriculture land which is situated close to a road or a residential area may be changed to a residential state, while a forest located in low and flat land with fertile soil may be converted to cultivation under food shortage pressure. Outputs from this block address the first research question.

In the econometric model the explanatory variables were regressed against the binary response surface of land use change during 2005-2010 (or the dependent variable). The binary response surface is the output of the overlaying operation between two satellite images (in this case, they were 2005 and 2010 images) where cells having different land use values from the two images were coded "1" and otherwise coded "0". The function resulting from regression was then transformed to produce the "probability of being changed" values (or so-called Y* values). It is a real number ranging from 0 to 1 for each PolyCell, where 1 indicates the highest probability of being changed. In other words, the "probability of being changed" value of a PolyCell shows how easily the land use of that PolyCell could be changed based on the history of land use change at the same cell with the preferable conditions which encourage the change to happen. A rich forest PolyCell which had been changed to medium forest between 2000 and 2005 could have a high probability of being changed to poor forest if it was located too close to a road network (the preferable condition). Section 5.1.2 provides more details on how Y* values were obtained.

In the Data Preparation block, historical remote sensing data and secondary data helped define decision making behaviour and "resistance to change" values (or Z values). The landscape itself may have a complex mechanism to resist change of land cover under different biotic and abiotic factors. The decision-making behaviour was then formulated as a set of decision protocols and transition rules for land use derived by analysing the spatial data and secondary data. Land use changes happen unevenly across landscape surfaces, more often and explicitly at some points but not at other points. The "resistance to change" value of a PolyCell on the landscape shows how difficult land use changes are through time. Z values were produced by ranking three binary response surfaces of 1995-2000, 2000-2005 and 2005-2010. It is a real number with value ranges from 0 to 1 to be compatible to compare with the Y* value of the same PolyCell. In other words, the Z value represents the probability of not being changed at a PolyCell of interest by looking at its history of land use changes from satellite images. A PolyCell that has not changed its land use during the three periods of 1995-2000, 2000-2005 and 2005-2010 would be more likely to remain stable in the next period and would have the highest Z value of 1. The algorithm for Z values will be presented in section 4.2. At this point, Y* values and Z values are comparable.

The Simulation block explains the mechanism of how the SeABM works. From this point, the second and third research questions were addressed. For the second research question, simulations were run without any policy interactions (the so-called baseline scenario). For the third research question, different policy alternatives (or scenarios) were simulated to reflect the outside intervention on land use change. In the core of the simulation, agents make their decisions about land use considering different driving factors and the resistance to change at any point of interest within the context of their household profiles. If all factors are preferable and they create a higher readiness to change value compared to resistance to change value, then the land use cover will be changed.

The simulation iterates every year from current year T_s to the last year T_e of a simulation (T_s and T_e are predefined by the user). All necessary data related to households (or agents) was extracted to the

PolyCell level. Each agent may manage several PolyCells when making decisions. At any year, T_j, the model will generate the "resistance to change" value (Z value) and the "probability of being changed" value (Y*) for all PolyCells, P_j, and compare them against each other. The SeABM also looks up possible land use change types and decision-making profiles to help agents decide whether to change or not. An agent's decision is posited to have a connection to its own world view and the labour capacity it has. Income sources, social networks and constraints are also factors affecting agent decisions. The details of potential land use change types and how a household profile influences land use change during the simulation process are explained from section 4.3 to 4.5. This process will be repeated until it reaches the predefined period of simulation. Data of preceding iterations are updated and used for succeeding iterations.

A nested loop using a Python script ensures that the simulation runs through every single PolyCell at each iteration. This demonstrates a tight integration of GIS and ABM in this SeABM simulation. At the end of the simulation, the land use status of a PolyCell, P_j, may change and dissolving PolyCells by land use allows geometric transformation. The power of GIS was deployed in this step to manipulate the transformation and then calculate and analyse the environmental outcomes, such as land use covers. Carbon sequestration was calculated through dependent carbon modelling which uses LULC types as inputs, while erosion will be evaluated through the LULC types and classes of slope. The analysis of forest fragmentation will be run on PolyCells having LULC classified as forest. Those are considered as environmental outcomes of the landscape.

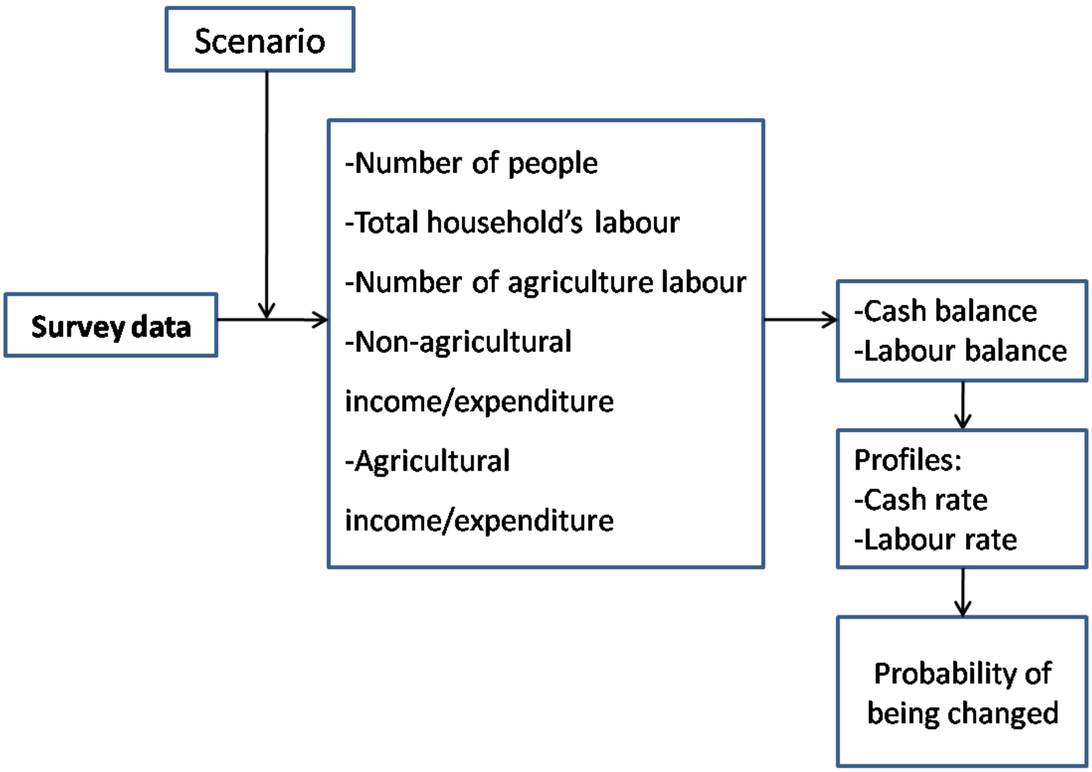
The structure of the SeABM used in this research will be open to revise parameters in order to test different scenarios, following the approach of Castella et al. (2005). The scenarios are designed to reflect different combinations of current and possible policies or development programmes in the study site. Comparing the results of simulations under scenarios shows the implications of those combinations. The design of scenarios will be explained in section 3.2.6.

3.2.5. Parameters and variables of SeABM

The projected period or cycle (T) of simulation is 10 years, starting from 2010 based on the availability of remote sensing data and demographic data. The source data file is a polygon feature class of land use in 2010. The details of how this feature class was constructed and the list of its data fields are described in spatial data collection and processing (section 3.2.3).

The simulation time step (t) is 1 year, meaning that the simulation will run repeatedly every year from 2010 until the iteration reaches 2020. At each iteration a set of variables (or attributes) needs to be calculated for each household (or agent). The simulation firstly calculates the household's demographic variables. The flow of calculation is described in Figure 15 below.

Figure 15. Calculation flow of household's demography



As can be seen in Figure 15, from the surveyed data (section 3.3.1) demographic variables like the number of people in a household, the availability of labour and labour dedicated to agriculture, income from both agriculture and non-agriculture sources, and the balance of labour and cash at the end of each year were calculated and their values later used to calculate household profiles, which will be mentioned in section 4.3 and section 4.5 when constructing the LULC trajectories. Each scenario designed for the simulation (section 3.2.6) will have different parameters such as growth rate of income, growth rate of population, amount of payment for forest environmental service and extra financial support that certain low-income households could get. Those parameters will be used as instructions for the calculations.

As documented in section 3.2.3, the spatially explicit Agent-based modelling in this research runs based on the PolyCells. Each record or row in the land use 2010 database (of the feature class) corresponds to a PolyCell. All variables are calculated for each row. Some of them have the same values in each row to represent the information for a household or at a household level. For example, all rows belonging to household A will have a value of M as it is the number of household members. Other variables may have different values in different rows because they carry the information at a PolyCell level or specific location, managed by household, on the land surface. For example, the values of land use type in rows may be different in reality. From the PolyCell level, the values for a household can be achieved by aggregating the values of all PolyCells (or rows in database) belonging to that household.

Depending on the output of the econometric model (section 5.1.1), some demographic variables will join with spatial variables to define the probability of the being changed in each PolyCell (or Y*'s value) and compare it to the resistance value (or Z's value) to determine whether a PolyCell is ready to change or not. If it is ready, then the corresponding land use conversion modules will be activated in simulation. How the land use conversion modules are set up is explained in section 4.4. In the simulation scripts, land use conversion modules are coded as functions running at the PolyCell level. Those functions have several inputs, including the current profile of the household which owns the PolyCell, current land use cover types of the PolyCell, elevation at the PolyCell's centroid, slope at the PolyCell's centroid, distance to river and road network from the PolyCell's centroid, and current year of simulation. The conversion module functions check the current profile of a household and activate all possible land use conversion modules. They change what land use cover types need to be considered and update those that need to be updated. In case the probability of being changed is lower than the threshold value, some land uses can still be changed. For instance, the forest cover types are updated, or perennial trees get older each year, as part of a natural process. It is worth noting again that within the 21 land use modules, some are activated at every year, while others can occur only every 5 years, while a few are available only after 10 years. The return of a land use conversion function is the land use cover type of a PolyCell.

The whole simulation process is the interaction between a Python script and polygon shapefile layer (containing PolyCells) as a data source. Attributes of the projected polygon shapefiles of the year 2020 (results of simulation with different scenarios) are grouped by household identification and land use cover types using the Dissolve function of ArcGIS 10.1. The output of this function is a new land use layer at the projected year for each scenario.

This simulation was designed to run with several scenarios as the impacts on the land use change. The simplest one is the baseline scenario or "business as usual scenario". The simulation with the baseline implies that there is not any kind of policy that intervenes with the land use change dynamics on the landscape. In other words, the baseline reflects the continuation of the current situation at the study area as closely as possible.

3.2.6. Scenarios for simulation

The last research question requires land use dynamics to be simulated under different policy interventions. The results of these scenarios determine how these external interventions affect the environmental outcomes of the landscape. GIS was deployed in handling, calculating, analysing and visualising the simulation results.

Policies are normally designed for higher administrative levels such as country, region, city/province or for collection of localities such as group of poor communes/districts. Policies influence the socioeconomic outputs through different means ranging from taxation to incentives and interest rates. In general, there are no specific policies, issued for communes like My Lam. However, based on central government policies many aids or programmes are flexibly interpreted for specific objectives at the local level to gain benefit from those policies. For example, poor households in rural areas may be given access to low or zero interest rate loans to promote husbandry or cash crops from the Vietnam Bank for Agriculture and Rural Development (Agribank). It is the only 100% state-owned commercial bank in Vietnam and has had a clear focus on agriculture since its establishment in 1988.This financial benefit is implemented as a part of the national poverty alleviation policies. Although this research uses the term "policy intervention" it needs to be clear that different policy scenarios are, in fact, different programmes that carry out the main spirit of different poverty alleviation and agricultural development related policies.

3.2.6.1. Local policies

Local authorities are given clear targets to promote the local economy by the central government. All post and pre-2010 policies for the study area were issued to support poverty alleviation and new rural development programmes, which were sourced from the upper levels of authority.

Based on the reports and planning documents for My Lam commune from 2011 to 2013, programmes were grouped into the following sets:

a) Poverty alleviation and promotion programmes for remote areas:

The aim of this group of policies is to reduce the poverty rate by supplying seedlings, fertilizers or cattle for poor households to create sustainable income.

b) Land use conversion and utilisation:

In this group, policies focus on introducing new crops or new technology to achieve the target in household income or household food consumption. This group includes several approaches such as:

+ Trials of new vegetables and medical plants in some small areas, these plants promise to bring higher income

+ Maximizing the areas for cereal to increase the average amount of cereal per head (up to 1,890 kg/person/year from 1,791/kg/person/year) to ensure food security

+ Converting the poor or exhausted forest areas (or classified as forest) to acacia hybrid or cashew to get higher value rather than leaving poor forests with very low chance of regeneration and restoration

+ Improving the old cashew areas by using a new variety which brings higher productivity

+ Protecting the forested areas to receive the payment for environmental services. Fees will be paid based on the size of protected area and forest quality. This is an annual payment for households

c) Social policies:

Social policies target several well-being indices which include:

+ Targeting the annual income of NZ$ 740 per person

+ Slowing the natural population growth rate of 2.7% down to 1.5%

3.2.6.2. Scenarios for policy intervention in 10 years

Based on information related to local policies and the structure of simulation framework, there were seven scenarios designed as below:

a) Baseline scenario:

This is a so-called "business as usual" scenario in which the population will grow at a rate of 2.7% per year as a normal rate. The number of mouths to feed in each household and the labours will also increase due to population growth. The cash balance of households may be deficit and lead to changes in their land uses to compensate for any loss. The land use changes will follow the historical trend described by several land use change modules results of analysing time-series satellite images. The output of the baseline scenario will be reported in section 5.2.

b) Low population growth scenario:

In this scenario the number of household members will be set to grow at a rate of 1.5% per year which reflects an attempt of population reduction policy. The decrease in population growth will also slow down the growth of labours and hence influence other demographic variable of households. This scenario is expected to drive the current land uses into less labour-intensive land uses due to the lack of labour forces.

c) High income scenario:

In this scenario, income growth rate was doubled from 5% (which was used for other scenarios) to 10% per year to meet the target of NZ$ 740 per head per year planned by local authorities. This scenario is expected to increase the balance of cash in each household which allows households to change less productive land uses into higher productivity land uses. For instance, households with spare money can invest into perennial crops by converting available suitable land uses.

d) Financial support scenario:

In this scenario, households with a cash balance lower than 0 can be subsidized a payment of NZ$ 300 as an incentive to help them have a chance to get a sustainable income. In the study area, poor households were provided with cattle or low interest loans to develop livelihood, other households who are not poor but in need of investment can search for a loan or borrow from other sources. It is difficult to integrate those details in a simulation, so it is set at an average of NZ$ 300 for the accessible finance for households having negative cash balance.

e) Cashew promotion scenario:

In this scenario, suitable areas for perennial crops are converted to cashew and existing cashew areas are maintained. Land use change modules number 4, 5, 6, 7, 8 are set up to change the current land uses, if eligible, to cashew after five years from start of simulation and land use change modules number 17, 18, 19 are also set to be changed to cashew but at the tenth year of simulation. Land use changes modules are analysed in section 4.4.

f) Acacia hybrid promotion scenario

This scenario is similar to the cashew promotion suitable areas for perennial crops are converted to acacia hybrid and existing acacia hybrid areas are kept. Land use change modules number 4, 5, 6, 7, 8 are set up to change the current land uses, if eligible, to acacia hybrid after five years from start of simulation and land use change modules number 17, 18, 19 are also set to be changed, if eligible, to acacia hybrid but at the tenth year of simulation.

g) Promotion of payment for forest environmental services (PFES) scenario:

This scenario assumes that a payment of NZ$ 25 will be paid for each hectare of protected forest regardless of the type of forest. This kind of promotion is expected to increase the quality and quantity of forest types intensively by upgrading forest type at every five years to a richer forest type. Related land use change modules are programmed to execute without any constraints in slope, elevation and proximity to roads and water. They are also allowed to bypass the comparison between probability to change (Y* values) and resistance to change (Z values). This scenario is expected to increase the forest area and quality creating a higher ecological value for the landscape.

Seven scenarios were synthesized based on the contents of different policies for the purposes of this research. Despite the attempt to make these scenarios as close to reality as possible, they do not necessarily reflect the opinion of the local authorities. The spatially explicit agent-based modelling was designed to test these scenarios to compare different effects of interventions on land use cover change in a 10-year period. The simulation framework was constructed to run with alterations of some variables to reflect the interventions from policies.

When testing the scenarios, some variables were changed and some variables such as agriculture investment and benefit were held unchanged. The purpose of using scenarios for this simulation is to analyse the differences of alternated input parameters in each scenario which may influence the output of simulation in different ways. In other words, it will help in understanding the different effects of policy alternatives on simulated land use. By using scenarios, this research will provide a tool to validate some policy interventions before they can actually be formulated.

3.2.7. Scripting for simulation

The simulation framework was built using Python scripting language, version 2.7. ArcGIS 10.1 gives a flexible way to automate data processing in general and geoprocessing in particular. The introduction of ArcPy module in ArcGIS 10.1 shows a tight integration of Python scripting in handling spatial data and promoting spatial related workflows in an ArcGIS environment. There are two ways to start Python scripts in an ArcGIS environment (from version 10). It can be run from an embedded Python window in

ArcGIS. Python interacts with spatial data in command line style or through an Integrated Development Environment (IDE) which gives more options to write, edit and run scripts.

The modelling for this research was designed to get inputs from the user, especially the scenarios. By using the parameters in the input, users can provide inputs to simulations to run specific data files with the desired scenario(s). The input data file for simulation is a feature class containing attribute data of 181 households at the year 2010 (or baseline year). Records (or rows) in the data table of this feature class are PolyCells, which were discussed in section 3.1.3.

In the main body of simulation scripts, several loops were used to go through records in the data table and several intermediate variables were calculated within the loops. The first loop starts from year 1 to year 10 (or from 2010 to 2020) as the simulation is designed for a 10-year period. In each year, the second loop will iterate through the 181 households in the data table of the feature class using their ID then calculate or update data in cells and allocate households into household profiles, where they are ready for being assigned land use conversion modules in land use conversion functions.

The scripts also calculated the Y* values and compared them to the Z values to activate the related conversion functions. The output shapefiles of simulations were dissolved by household ID and land use types. The script mainly uses the Arcpy module to handle the attribute data in table of the shapefile using SearchCursor and UpdateCursor and different geoprocessing and mapping tasks. However, it also calls other external modules such as operation system module ("os") to manipulate files and scientific module ("scipy") to calculate the values cumulative distribution function (or Y* values). A full version of working scripts is provided in Appendix 3 for reference.

3.3. Data collection and processing

Based on the proposed methodology and approach for this research, many types of data were required. The interdisciplinary approach requires both demographic and spatial data. This section explains how data was collected, managed and prepared for different parts of this research, from overall socio-economic analyses to econometric model and SeABM simulation. A survey for demographic data and secondary data as well as collection and processing of spatial data layers and their transformation will be detailed in this section.

3.3.1. Survey

This section explains how demographic data were collected for this research using a survey questionnaire at the selected study area. After selecting My Lam commune as the study area that satisfied all criteria of this research, the initial contact and secondary data collection were made. A survey questionnaire was designed and implemented to obtain basic household data. Other necessary and available data such as paper maps, electronic maps, documents issued by the local authorities and commune land allocation sketches were also collected.

3.3.1.1. Questionnaire design

The questionnaire was designed to collect demographic, socioeconomic and land use information from households in My Lam commune. There were nine main questions with several sub questions fitted on 3 sides of A4 paper. The questionnaire was required to be short and easy to understand and answer in a duration of 20 to 30 minutes to reduce disturbance for interviewees.

There were many questionnaire design techniques deployed ranging from specific answers to questions using the Likert scale, weighting and open-ended questions (Babbie 2005). Details of the questionnaire can be found in Appendix 1. The first group of questions (1.1 to 1.4) focuses on household demographics, household composition and migration status. The second group (2.1 to 2.2) investigates the income and expenditure information of each household. The next group of questions (3.1 to 3.3) checks the land use decision making history with its associated factors and asks the interviewee to rank them. Question 3.4 in this group asks interviewees to list all current land plots managed by each household, their tenures, estimated income or cost associated with production on land, and their historical status of land use. Question number 4 explores intentions to change land use in each interviewee's opinion. Photos representing good agricultural practices were shown to interviewees to see if they want to change their land use to follow those good practices. Exemplars (on photos) were productive rice field, productive maize field, coffee, pepper and rubber plantations, poor forest for husbandry, medium forest and rich forest. The photos are included in Appendix 2. Question 5 requires interviewees to weight the seven factors land use decision making is believed to be based on. They are financial factors, the availability of labour in the household, the price of common commodities (rice, corn, coffee, cashew and timber), proximity to water, road or residential areas, the influence of their neighbours on decisions, and slope and elevation factors. Factor weights varied from 1 to 10 for each factor and the total weight of seven factors must be 10. The next group of three questions (6, 7, and 8) explored the opinion of interviewees regarding the conversion of forest land to other purposes, what would the difficulties and economic incentives be. In the last group, question number 9, interviewees were asked to describe their perception of distance, slope and elevation. For example, respondents were asked to describe what distance they think is far and what elevation they think is high.

By answering the questionnaire, interviewees were assumed to provide enough demographic information to combine with the spatial data in the econometric model. The questionnaire also aimed to capture the dynamics of land use in each household based on the description from the land users. This is an important set of information needed to understand the history of land use at the local level, the preferences of land use conversion and temporal pace of land use changes. By weighting different potentially influencing factors on decision making processes, interviewees revealed their consideration of different factors in their final land use decision. This information helped inform this research on the behaviour of the decision makers. Other information regarding the opinion of interviewees on forest conversion helps in understanding the historical and future trends of forest degradation as well as the "willingness to accept" cost needed to maintain forest. For further analyses of the trajectories of land use change, perception about distance, slope and elevation from interviewees were very useful. The land use cover at each point of interest on the landscape surface was extracted from temporal satellite images and then associated with the distance and slope data to give more understanding beyond the land use changes through time.

3.3.1.2. Interviews

The plan and budget for the survey in My Lam commune were drafted in January 2013 and finalised at the end of February 2013. The questionnaire, information sheet and consent form were also finished by the end of February 2013. The application to the Human Ethics Committee of Lincoln University was lodged on 1st Mar 2013 and approved on 2^{sd} May 2013. All documents were then translated into Vietnamese for use at the study area.

The survey was conducted between 4th and 16th June 2013 with permission of the People's Committee of My Lam commune. The local authority also granted access to the household data in My Lam commune to prepare for sampling for the survey. All households who have farm land would be approached to ask for participation in the survey. Criteria for choosing participants for this research were designed to acquire the appropriate information without violating any human rights. All household heads with farmland in the study area were potential interviewees.

Estimating the smallest sample size for this research is complicated. However, the outputs of the survey were the inputs for the econometric model using logistic regression, so the estimation of sample size was adopted from other authors' experience. Peduzzi et al. (1996) suggested the following formula:

$\mathbf{N} = 10*\mathbf{k}/\mathbf{p}$

Where **p** is the smallest proportion of negative or positive cases in the **N** population and **k** is number of dependent variables (or covariates). In this research, three dependent variables (**k** = 3) were deployed and if proportion of negative cases (or land uses that were not changed) is 50% (**p** = 0.5), then the least number of samples is:

$\mathbf{N} = 10*3/0.5 = 60$

However, it was suggested by Long (1997) that if the minimum of samples is less than 100 then it should be increased to 100.

From the list and associated information of more than 200 households provided by local authority, a sub set of 181 households was selected. Participants who were under 16-years-old and those not involved in any type of farming activities were excluded. The sample number was maximized to make sure it was distributed across the study area, as the locations matter in this research. Heads of households (could be male or female) were the primary respondents. It was deemed acceptable if they wanted to consult other adult members in their households who were involved in decision making to respond to questions.

The aims and the content of the interviews were introduced through trusted local people in the community to make sure that all households in the study area knew about the project and the proposed interview using printed information sheets. In this research, one officer working for the local authority supported the dissemination of information to four hamlets in the study area. Potential participants were then personally contacted and asked for voluntary participation. Once the purpose of the survey and the content of the questionnaire were made clear to participants, 181 participants agreed to give information. A convenient time and date were set for voluntary participants to have interviews. Before giving any information, interviewees were asked to read and sign the consent form in Vietnamese to fully understand his/her right to withdraw from the interview anytime he/she wanted.

The survey was carried out in June 2013 when local people were busy starting the new paddy rice season. For this reason, the questionnaire was designed to be carried out in 20 - 30 minutes to avoid disruption to the interviewees. Interviews were carried out between 7am and 8pm at interviewees' houses or at the fields where they worked. In each interview, the aims of the interview were discussed, and interviewees were asked to sign the consent forms. Then he/she was asked to provide some basic demographic information, current and historical land use decisions, and factors that have affected these decisions. Eight photos of good practices were showed to help interviewees to identify preferences for land use on specific types of land under their management. Interviewees were also asked to give information about preferred location and elevation for different land uses.

To support interviewees to understand the weighting of several factors for question number 5, interviewees were provided with 20 small stones and seven cups, each cup was tagged with a factor and interviewees were asked to put more stones in factors (or cups) that they thought more important. The more important the factor (based on each interviewee's opinion), the more stones that cup got. One stone is equivalent to 5% and 20 stones are equivalent to 100%. Interviewees were given time to be sure about their decision before their answer was recorded. Use of the stones and cups appeared to help interviewees understand the weighting system and assured that responses closely reflected the interviewees' beliefs. Interview information was used to develop the econometric model that predicted land use changes based on a range of demographic and geographical estimators.

Demographic data from this survey were processed and used as attribute data for PolyCell as mentioned in section 3.2.3.

3.3.1.3. Demographic data entry and management

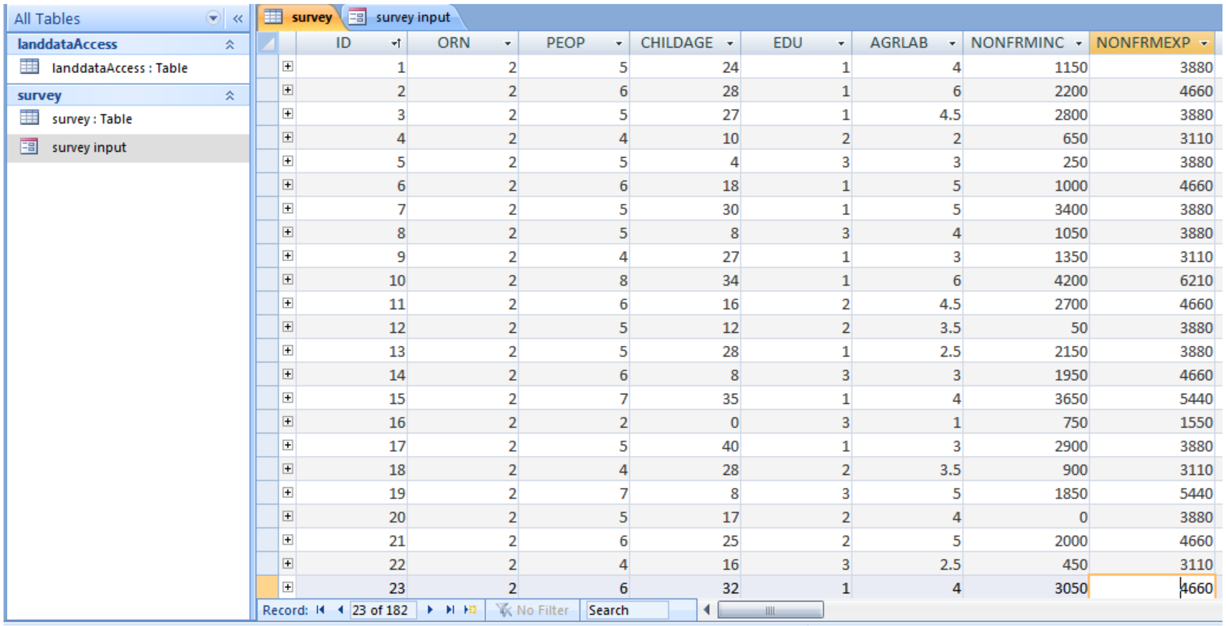
The consistency of completed questionnaires was checked every day during the survey to assure quality of information. When any mistakes were identified, they were corrected before returning to New Zealand by asking the interviewees to clarify their responses. Since the questionnaire form was short and the questions were quite understandable for most of the interviewees, no significant errors were found.

Based on the questionnaire, a Microsoft Access data input form was designed for data entry. The form was designed based on the questionnaire layout, allowing easy tracking and checking of input errors. Data were stored directly in a Microsoft Access database and could be easily be exported to Microsoft Excel 2007 and Stata10 (StataCorp 2009) for statistical analysis.

For convenience, each questionnaire was assigned a unique ID (identification) number. This is also the ID number for each household. This ID number is a key field in a Microsoft Access 2007 database. The list of households and their matching IDs was kept separately to ensure the anonymity of each household.

A snapshot in Figure 16 represents the basic structure of the database designed for this research.

Figure 16. Microsoft Access 2007 database designed for this research



ID	ORN	PEOP	CHILDAGE	EDU	AGRLAB	NONFRMINC	NONFRMEXP
1	2	5	24	1	4	1150	3880
2	2	6	28	1	6	2200	4660
3	2	5	27	1	4.5	2800	3880
4	2	4	10	2	2	650	3110
5	2	5	4	3	3	250	3880
6	2	6	18	1	5	1000	4660
7	2	5	30	1	5	3400	3880
8	2	5	8	3	4	1050	3880
9	2	4	27	1	3	1350	3110
10	2	8	34	1	6	4200	6210
11	2	6	16	2	4.5	2700	4660
12	2	5	12	2	3.5	50	3880
13	2	5	28	1	2.5	2150	3880
14	2	6	8	3	3	1950	4660
15	2	7	35	1	4	3650	5440
16	2	2	0	3	1	750	1550
17	2	5	40	1	3	2900	3880
18	2	4	28	2	3.5	900	3110
19	2	7	8	3	5	1850	5440
20	2	5	17	2	4	0	3880
21	2	6	25	2	5	2000	4660
22	2	4	16	3	2.5	450	3110
23	2	6	32	1	4	3050	4660

As seen in Figure 16, the database was encrypted so it could be accessed only by the student and supervisors. The hard copy version of the filled questionnaires was kept in a secure place.

3.3.2. Secondary data

Land use data were collected from the Provincial General Statistics Office (PGSO). Land use classifications, their corresponding area and land use dynamics for several years, depending on their availability, were extracted for use in further analyses in this research. Secondary data were also collected in discussion with local community members and observation during the fieldtrip. Information about migration, population and agricultural practices were discussed with people living in the study area. This information helped with understanding the context of development in the commune and to shape the general idea of how local people were involved in different activities and made land use decisions.

Some information was taken from the annual socioeconomic reports issued by local authorities. Information from those reports was used to design the scenarios for later simulations (discussed in section 3.2.6). Employees working for local authorities and experts from the National Institute of Agricultural Planning and Projection (NIAPP) were also consulted during the data collection and analysis phase. These sources contributed valuable information related to soil fertility, suitable crops and plantations for the local area, habits of farmers in agricultural activities and history of land use.

3.3.3. Spatial data collection and processing

Depending on their availability and accessibility, several layers of spatial data were collected and processed for this research.

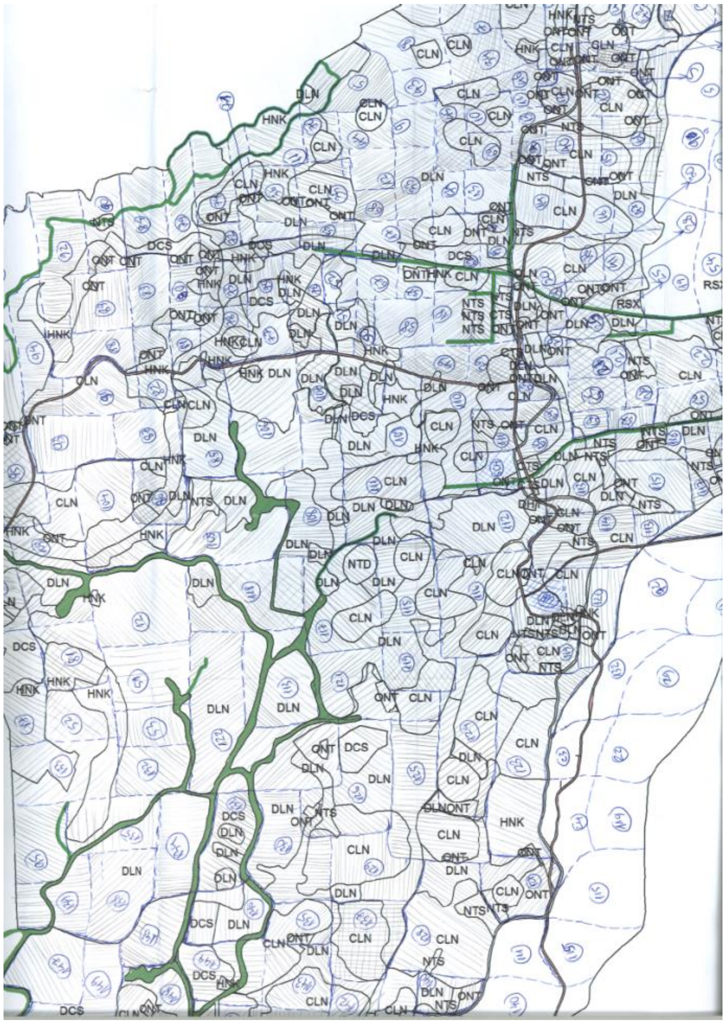
The projected WGS 84 zone 48N coordinate system was used for this research. This helped to decrease the effect of distortion in data analysis. The unit of length is metres and unit of area is hectares. Spatial materials having different coordinate systems were reprojected to WGS 84 zone 48N to ensure a common coordinate system.

All spatial data were stored and processed in a personal geodatabase in ArcGIS named *dataprocessing*. This is the main and only spatial database used in this research.

3.3.3.1. Sketch map of land use by households

After answering the questionnaire, interviewees were asked to identify the boundary of their land use on the provided sketch map. This simple map has the commune boundary, hydrology network, road network and different land marks to help the interviewees easily locate their land parcels. It was printed on A0 format with dimension of 84.1 by 118.9 centimetres. A fraction of this sketch map was scanned and represented in Figure 17.

Figure 17. Sketch map of land use allocation in My Lam commune



This activity was also supported by a local employee who well understood land use and land allocation at the study area. The sketch map was scanned and represented as seen in Figure 17 and was used to analyse the spatial distribution of land parcels in the study area. It was digitised into ArcGIS to serve further analyses and modelling.

It is worth noting that this sketch map was built up by hand drawing, to approximately represent the land use allocation for each household in the study area, based on what households estimated, and was not the product of a cadastral survey. Consequently, the boundary, shape and location are not highly precise. However, this does not affect the effort to build up a SeABM in this research. The sketch map was cut into 11 small pieces and scanned into JPEG files. Those images were then rectified using a set of control points using the Georeference tool in ArcGIS. By cutting sketch map into small pieces, it helped to scan them easily with an A3 scanner to minimise distortion.

The 11 scanned and rectified images were then mosaicked together in ArcGIS to build the whole scanned sketch map. This scanned sketch map was then used as an underlay for digitizing the land use boundary of each household. ArcGIS 10.1 provided a convenient environment with supportive tools for digitising and creating associated attribute data for polygon feature class (*dataprocessing/ML_land_by_ID*).

In the associated attributes, digitised polygons carried IDs which are identical to the IDs of the 181 households participating in the survey. Those IDs will be used as link keys to merge data from different sources. The feature class *ML_land_by_ID* held several multipart polygons as a result of the digitising process. Those polygons were then converted into single part polygons by using the Explode polygon tool in the Advanced Editing toolset of ArcGIS (*dataprocessing/ML_land_by_ID_single)*. The aim of archiving single part polygons in a feature class is for further analyses which need the centroids (or geographical centres) for data extraction and processing.

The Topology tool was also run to validate the single part polygon feature class, to ensure that it did not contain overlapping polygons or unwanted slivers. To update the land use type of each polygon in *ML_land_by_ID_single* the land use cover image from 2010 (*dataprocessing/Landcover_MyLam_2010.img*) was converted into vector (*dataprocessing/ ML_landcover2010_img*) and then land use data were taken by using Intersect tool in the Geoprocessing toolset of ArcGIS. The 2010 image was considered as the most recent land use data able to be obtained for this research. The intersect process computed a geometric intersection of the two input feature classes. Features or portions of features which overlap in all feature classes were written to the output feature class *dataprocessing/ML_land_intersect_2010*.

3.3.3.2. Satellite images

Processed satellite images of land cover were supplied by a contact from SNV. They were clipped using a 500m buffer around the My Lam boundary (*dataprocessing/ML_500mbuf*). This process made the underlying raster extent larger than the commune's boundary to avoid Nodata data type from the points lying on that boundary when using the Spatial Analyst tool "Extract Values to Points”.

All images were rectified to the current projected coordinate system and stored in the geodatabase with following names:

Landcover_MyLam_1995.img

Landcover_MyLam_2000.img

Landcover_MyLam_2005.img

Landcover_MyLam_2010.img

where suffixes of 1995, 2000, 2005, 2010 are the years of images.

These images had a spatial resolution of 20 meters and they were used for many purposes, from extracting values for modelling and analyses to overlaying and calculating to create new rasters. By using these images, the land use trajectories were investigated (as in section 4.3) and the dynamic of land use cover changes throughout the years have been explored. These data are very important for further analyses in this research.

3.3.3.3. Vector data

A vector dataset of land use of the My Lam commune was provided by a contact from the Central division of National Institute of Agricultural Planning and Projection (NIAPP). This dataset was created in MapInfo (.TAB format) based on the land use survey in 2010. It had boundary of the My Lam commune, contour lines, elevation sample points, hydrology network, road network, current land use status and labelling annotation. From this dataset, the commune boundary, land use status polygons, hydrology network (lines, polygons) and contours were extracted and converted to feature classes in the *dataprocessing* geodatabase. Those vector data will further be processed (section 3.2.3.4) to achieve different secondary spatial data for this research.

3.3.3.4. Processed spatial data

Based on the satellite images and vector data features collected from sources, other intermediate spatial data were generated to serve in other analysing steps in this research. Data generation and creation were carried out in the environment of ArcGIS with its several useful tools. A fishnet (a net of rectangular polygons) with dimensions of 20m x 20m was created using XTools Pro - an extension for ArcGIS (Data East Soft, 2015). This polygon feature class has the spatial extent of *ML_500mbuf* feature class and is stored in the *dataprocessing/ML_500mbuf_fishnet*feature class. This fishnet was used to create the PolyCells in this research. The method and the advantages of using PolyCell were presented in section 3.2.3.

A feature class of contours (*dataprocessing/contours*) was used as the input for the Spatial Analyst tool called "Topo to Raster" to interpolate a hydrologically correct raster surface from line data. The output of this tool was the digital elevation model (DEM) of the My Lam commune with a better resolution than the 90m resolution of the DEM taken by Shuttle Radar Topography Mission (SRTM). The interpolated DEM is stored in *dataprocessing/DEM* and has a minimum elevation of - 17.87m.a.s.l and a maximum value of 362.83m.a.s.l. This raster will be used to extract the elevation values for the econometric model in section 3.1. Slope and hillshade surfaces were derived using the DEM as an input. They were stored at *dataprocessing/slope* and *dataprocessing/hillshade*. Both elevation data and slope data will also be used as estimators in the econometric model in section 3.1.

The hydrology network (*dataprocessing/hydrline* for line objects and *dataprocessing/hydrpol* for polygon objects) was converted into a raster layer and the "Euclidean Distance tool" (Spatial Analyst) was launched to calculate the Euclidean distance surface, which indicates the distance from any point of interest to the hydrology network. This surface is located at *dataprocessing/ML_dist_wtrsource*. Similar steps were used to generate the Euclidean distance surfaces of the road network and residential areas.

The 'Focal statistic variety" of ArcGIS was deployed to calculate the number of land uses (or variety of land uses) in a rectangular (neighbourhood) of 50mx50m of each location of interest. The inputs of this tool were the land use cover satellite image of 1995, 2000 and 2005 and the outputs were the focal rasters for 1995, 2000 and 2005. The "Focal statistic variety" exhibited the diversity of land uses in a pre-set area around the location of interest. These rasters were then served as estimators in the econometric model in section 3.1.

3.4. Probit regression

3.4.1. Sampling method for input data of the econometric model

The hypothesis of this research was that land use cover changes are influenced by the combination of human induced factors (or demographic factors) and spatial related factors. To prepare the input data for an econometric modelling approach, both demographic and spatial data were gathered. The integration of spatially explicit data in an econometric model demonstrated an interesting approach in this case. The demographic data were collected through the on-site survey using a questionnaire, as outlined in section 3.3.1.2. The spatial data, with two main formats being vector and raster data, were collected and processed using ArcGIS 10.1. The important spatial data for the econometric modelling are land use cover, elevation and slope.

Since the econometric model runs on Stata, it does not directly handle the spatial data from ArcGIS, and those spatial data need to be converted into a supported format. There were 181 households (or 181 records) in the demographic database (in Microsoft Access format) so the spatial data needed to be extracted to synchronize with those 181 records. Households' IDs were used to link data from different sources.

The GIS allows the household-related spatial data to be exported to a representative point for each household. In the spatial database, a point (or location with exact longitude and latitude) is placed within the area (polygon or land plot) managed by each household. However, a household could have more than one land plot or polygon, so the sample points were not randomly generated in each polygon or centroids of polygons but followed a special design rule.

A feature class of land use data in 2010 (*dataprocessing/ML_landuse_intersect_2010*) was used to generate the sample points. It holds 2170 polygons belonging to 181 households. The key task here was how to select a representative polygon among many polygons for each household. Firstly, the resistance values (Z values) from the resistance surface (to be discussed in section 4.2.) were extracted to the centroids of all polygons of *ML_landuse_intersect_2010* feature class. Then, based on those values, in each household, one polygon with the lowest resistance value was selected. It was assumed that the locations showing low resistance to change could be changed more easily.

The identification between polygons and their centroids was used to identify the corresponding polygons and at the end of this step there were 181 polygons representing the 181 households. This subset of 181 polygons was then used as the target to generate 181 random points; one was randomly drawn within each polygon with the support of ArcGIS. That set of sample points satisfied several conditions: (i) each random point represented a household, and (ii) they were drawn randomly (within the limited space of the 181 polygons) to ensure randomness for data collection and extraction, but (iii) were localized in the polygons having the highest possibility of land use cover changes. Output point feature class was stored at *dataprocessing/ML_land_intersect_2010_centrd_181lowthres.*

Locations of those 181 points of low resistance values with their other spatial characteristics were assumed to have a strong correlation to the change of land use cover.

All necessary spatial data were extracted using the 181 sample points and the final table of attributes was reformatted and merged with the demographic dataset to be fed into the econometric model and analysed with Stata10 (to be presented in section 5.1.1).

3.4.2. Probit regression estimation

The choice of probit regression as the econometric model for this research was discussed in section 3.1.

Based on the collection of satellite images, those from 2005 and 2010 were the latest available data close to the period when demographic data was collected in 2013. The trend of land use cover change during the2005-2010 period was assumed to have a strong correlation to the demographic data in 2013.

The binary surface for 2005-2010 was produced by overlaying the 2005 image with the 2010 image. Pixels having different land use codes between 2005 and 2010 carried a value of "1", otherwise, they carried a value of "0". The 181 sample points (generated in section 3.4.1) were used to extract raster values for the econometric model, used in this case to extract the 2005-2010 land use change dataset. This dataset was chosen as the dependent variable with the name of **chg0510**.

By using the same method as above, the 2000-2005 land use change dataset was extracted and used as an independent variable under the name **chg0005**. The potential independent variables for the regression are listed in Table 1 below.

Table 1. Variables for the econometric model

Variable	Type	Description
chg0510	Binary, Dependent, spatial data	Indicates the status of land use cover change at the same location from 2005 to 2010. Values were extracted from overlaying satellite images in ArcGIS.
chg0005	Binary, Independent, spatial data	Indicates the status of land use cover change at the same location from 2000 to 2005. Values were extracted from overlaying satellite images in ArcGIS.
elv	Continuous, Independent, spatial data	Indicates the elevation at a point of interest, measured in metres. Negative values indicate the depth under water. Those values were extracted from the elevation raster using the 181 sample points. The elevation raster was generated using the contours of My Lam commune. (section 3.2.3.4)

Variable	Type	Description
slp	Continuous, Independent, spatial data	Indicates the slope value at a point of interest, measured in degrees. Values were calculated from elevation raster using ArcGIS (section 3.2.3.4).
disbuild	Continuous, Independent, spatial data	Indicates the Euclidean distance from a point of interest to built-up areas. The Euclidean surface was calculated based on the infrastructure data from 2010 (section 3.2.3.4).
diswtr	Continuous, Independent, spatial data	Indicates the Euclidean distance from a point of interest to hydrology or water surfaces. The Euclidean surface was calculated based on the infrastructure data from 2010 (section 3.2.3.4).
distrans	Continuous, Independent, spatial data	Indicates the Euclidean distance from a point of interest to a major transportation network. The Euclidean surface was calculated based on the infrastructure data from 2010 (section 3.2.3.4).
focal95	Continuous, Independent, spatial data	Indicates the number of land use (variety) in a rectangular (neighbourhood) of 50mx50m of each location of interest. Those values were obtained by deploying the Focal statistic variety on land use cover satellite image of 1995 in ArcGIS environment (section 3.2.3.4).
focal00	Continuous, Independent, spatial data	Indicates the number of land use (variety) in a rectangular (neighbourhood) of 50mx50m of each location of interest. Those values were obtained by deploying the Focal statistic variety on land use cover satellite image of 2000 in ArcGIS environment (section 3.2.3.4).

Variable	Type	Description
focal05	Continuous, Independent, spatial data	Indicates the number of land use (variety) in a rectangular (neighbourhood) of 50x50m of each location of interest. Those values were obtained by deploying the Focal statistic variety on land use cover satellite image of 2005 in ArcGIS environment (section 3.2.3.4).
agrlab	Discrete, Independent, demographic data	Indicates the number of the labours that were involved in agricultural activities in each household. Those values were collected from the survey. Agriculture-involved labour is less or equal to the number of people in each household (section 3.2.1.4).
farmprf	Discrete, Independent, demographic data	Indicates the profit of each household gained from agricultural activities. Farm profit was calculated as the net of farm income deducted from farm investment. Farm income and investment were allocated based on the area managed by each household and the basic cost and income collected by interviewing local people on a fieldtrip (section 3.2.1.4).
farminvs	Discrete, Independent, demographic data	Indicates the investment of each household into agricultural activities (section 3.2.1.4).
nonfarmprf	Discrete, Independent, demographic data	Indicates the net income of each household from non-agricultural activities (section 3.2.1.4).
balc	Discrete, Independent, demographic data	Indicates the cash balance of each household after deducting all kinds of expenditure from the total income of both agriculture and non-agricultural activities. In that calculation, non-agriculture income and expenditure were collected from the survey and rounded up to a convenient level (section 3.2.1.4).

The proposed variables in Table 1 include both demographic and spatially related variables. The probit regression predicted the probability of land use cover change in 2005-2010 based on those variables. The syntax of probit regression in STATA10 is:

```
probit  dependent_var independent_var(1) independent_var(2)...independent_var(n)
```

The model estimation was carried out by running several probit regressions against different combinations of variables until a statistically robust model was achieved. The best regression model that could be achieved during the estimation had the land use cover changes over the last 5-year period (or **chg0510**) as dependent variables and the land use change from 2000 to 2005 (**chg0005**), elevation (**elv**), and number of agricultural labours of each household (**agrlab**) as independent variables.

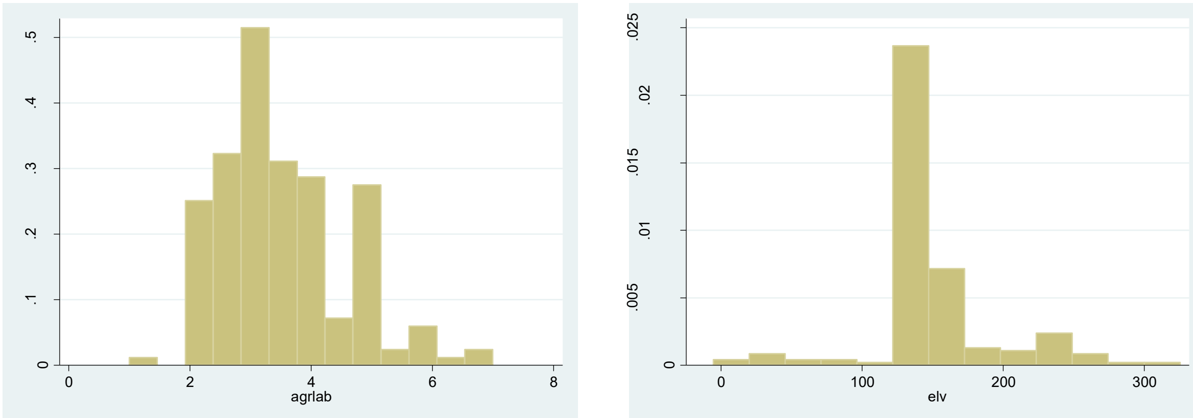
Table 2. Summary for input variables of probit regression

```
. sum  chg0510 chg0005 elv agrlab
```

Variable	Obs	Mean	Std. Dev.	Min	Max
chg0510	181	.3093923	.4635253	0	1
chg0005	181	.7679558	.423308	0	1
elv	181	147.0009	44.26254	-5.63732	325.282
agrlab	181	3.480663	1.124679	1	7

As presented in Table 2 all potential variables for probit regression have the full range of data with 181 observations; no blank or “Nodata” data type instances were detected. **chg0005** is the same type of binary variable as the dependent variable. **elv** ranges from -5.637m to a 325.282m. The negative elevation shows the deep location compared to sea level while the positive value of 325.282m is a moderate height compared to sea level. The availability of agricultural labour in each household varies from 1 to 7 labourers. Since both **elv** and **agrlab** are continuous variables, their histograms were drawn and presented in Figure 18.

Figure 18. Histograms of agricultural labour and elevation



Both histograms in Figure 18 show some similarity to the normal distribution. However, observations seemed to be distributed more in the right tail in both cases.

A summary for the dependent variable (**chg0510**) is represented in Table 3 below.

Table 3. Distribution of dependent variable

chg0510	Freq.	Percent	Cum.
0	125	69.06	69.06
1	56	30.94	100.00
Total	181	100.00	

Table 3 shows that there are no absent or "Nodata" values for the dependent variable to be eliminated. The distribution of observations is 69.06% for 0 value (no change) and 30.94 % for the 1 value (change). The result of probit regression will be reported in section 5.1.1. It defined the driving factors which affected the land use changes.

3.5. Estimation of environmental degradation

As summarised in section 1.1 the environmental degradation is diverse, ranging from landslide, erosion to draught. They are the consequences of bad management and over exploitation of natural resources. To quantify or qualify all environmental degradation it needs several approaches. For this research scope, there are some kinds of environmental degradation such as carbon sequestration, soil erosion and vegetation fragmentation which can be estimated based on the land use cover outputs of the SeABM simulation.

3.5.1. Carbon sequestration

Carbon content calculated based on the vegetative covers could serve as an indicator for environmental quality. The carbon sequestration is estimated based on the data from Table 4.These data are adapted from a recent report of a SNV's REDD+ project in Lam Dong province (Gibbon et al. 2011). It is worth noting that this is just an assumption about carbon sequestration adapted from other research. There was no ground truth at the study area for these numbers in Table 4.

Table 4. Carbon content of different LULC types

Land use land cover types	Carbon (tonnes/ha)
Medium forest (standing volume: 80-150m3/ha)	103
Poor forest (standing volume: <80m3/ha)	87
Mixed forest	87
Bamboo forest	87
Young forest (regeneration)	35
Bare land	4
Rice field	5
Other annual crops (cassava, maize)	5
Forest plantations (acacia hybrid)	33
Cashews	50

Source: Ogonowski and Enright, 2013

According to Hoang et al. (2010), the carbon content of all LULC categories that are not natural forests should be set to zero. However, Ogonowski and Enright (2013) in their report argued that the carbon content of some crops can be substantial, for instance the case of forest plantations and rubber plantations. That is why LULC types differentiated from natural forests have their carbon contents represented in Table 4. The carbon sequestration was calculated for each scenario based on the area of each land use type in the simulation output for that scenario. The result will be reported in sections 5.2 and 5.3.

3.5.2. Soil erosion

LULC change is one of the negative impacts affecting the speed of soil erosion. Losses of above ground covers stimulate the losses of top soil layer being washed by rain and wind. Those losses degrade soil quality and soil structure. The speed of erosion depends on the steepness, the characteristics of the LULC and physical characteristics of soil. Basic data for erosion in Lam Dong province were extracted and adapted from research by Hai and Dung (2014).

The relation between soil erosion and classes of slope is presented in Table 5. The percent rise slope data for the study area are broken into five classes which all follow a basic principle: the steeper the slope is, the higher the soil erosion is. Forest cover has a lower soil erosion compared to other land use cover types. Both annual and perennial crops have reasonable soil erosion on high slopes.

Table 5. Average speed of erosion by types of LULC and slopes in 50 years

Type of LULC	Class	Slopes (degrees)	Slopes (%)	Erosion (t/ha/y)
Natural forest (Medium forest, Mix forest, Bamboo forest, Poor forest, Young forest)	1	0-5	0-8.75	0.5
	2	5-15	8.75-26.79	1.75
	3	15-25	26.79-46.33	2.1
	4	25-35	46.33-70.02	3.8
	5	>35	>70.02	5.05
Plantation forest (Acacia hybrid)	1	0-5	0-8.75	1
	2	5-15	8.75-26.79	2.1
	3	15-25	26.79-46.33	6
	4	25-35	46.33-70.02	7.5
	5	>35	>70.02	8
Cashew	1	0-5	0-8.75	16*
	2	5-15	8.75-26.79	19.8
	3	15-25	26.79-46.33	29
	4	25-35	46.33-70.02	33*
	5	>35	>70.02	20*
Annual crops (cassava, maize)	1	0-5	0-8.75	10
	2	5-15	8.75-26.79	17
	3	15-25	26.79-46.33	29
	4	25-35	46.33-70.02	34
	5	>35	>70.02	34*

*Source: Adapted from Hai and Dung, 2014 * - interpolated value from data in source to fill the blank data*

Some missing values in Table 5 were interpolated based on the information from Hai and Dung (2014). The models predicting erosion of different vegetative covers have been used to interpolate the values of some slope categories presented in the study area.

Based on the classes of slope and the areas of land uses from the result of simulation for each scenario, the average speed of erosion in the study area was calculated and is presented in sections 5.2 and 5.3.

3.5.3. Vegetation fragmentation

Forest fragmentation caused by breaking up large and entire forested areas into smaller pieces or patches is one of the major threats to the conservation of biodiversity and ecological functions of forests (Harris 1984, Terborgh 1989, Forman 1995, Rochelle et al. 1999, Loyn and McAlpine 2001). Forest fragmentation isolates the habitats for wildlife and causes erosion and water contamination due to increasing water runs surface. According to Saunders et al. (1991), Forman (1995) and Haila (1999), forest fragmentation can be summarised into three categories: reduction in area (or size) of forest patches; the isolation level and the loss of overall connectivity; edge effect and disturbance from surroundings. Causes of forest fragmentation can be both natural processes and human land use activities (Harris 1984). In the study area, human-induced forest degradation seems to be the predominant cause.

FRAGSTATS - Spatial Pattern Analysis Program for Categorical Maps (McGarigal et al. 2009) became popular in the fragmentation research community. It helps to compare the fragmentation of different landscapes, or different moments of the same landscape can be compared against each other to assess the diversity level.

Patch Analyst is a successor to and more flexible than the original FRAGSTATS program in terms of directly processing the polygon data in ArcGIS environment (Rempell et al. 2012). It reports several different landscape metrics for both landscape and class level, which were defined by McGarigal and Marks (1995). Classes are forest covers or land uses types present on the landscape. The landscape may have more than one class on it.

Patch Analyst uses polygons of land uses as inputs and offers 44 landscape metrics. However, for current research only 8 metrics were selected to analyse. They are defined as below with their abbreviations in Table 6 below.

Table 6. Landscape metrics

Metric	Description
Area-weight mean shape index (AWMSI)	If this metric is greater than 1, patches in all classes have irregular shapes. If it is equal to 1, patches have similar shapes to circular or rectangular shapes, which have an index of 1. The more irregular the shapes are, the stronger the likelihood of being naturally formed rather than being managed by humans.
Edge density (ED)	Edge density equals the sum of the lengths (m) of all edge segments involving the corresponding patch type per hectare. It reports edge length on a per unit area basis that facilitates comparison among landscapes of varying size.
Mean Patch Size (MPS)	This is the average patch size. At class level, it is the average size of every single class; at landscape level, it is the average size of all patches in the landscape.
Number of Patches (NumP)	This metric is the total number of patches in the landscape if analysing at landscape level, or Number of Patches for each individual class if analysing at class level.
Patch size coefficient of variation (AREA_CV)	This is preferable to standard deviation for comparing variability among landscapes. Patch size coefficient of variation measures relative variability about the mean (i.e. variability as a percentage of the mean), not absolute variability. Thus, it is not necessary to know mean patch size to interpret the coefficient of variation. Nevertheless, patch size coefficient of variation also can be misleading with regards to landscape structure in the absence of information on the number of patches or patch density and other structural characteristics.
Patch size standard deviation (AREA_SD)	This is a measure of absolute variation; it is a function of the mean patch size and the difference in patch size among patches. Thus, although patch size standard deviation conveys information about patch size variability, it is a difficult parameter to interpret without doing so in conjunction with mean patch size because the absolute variation is dependent on mean patch size.
Class area (CA)	This measures the area of a particular patch type of landscape. This parameter is important because many species may require a certain size of suitable habitat (or particular type of forest) to be present.

Those metrics as in Table 6 will be used to assess the fragmentation of simulated landscapes by analysing the forest patches in those landscapes. Each simulated landscape in sections 5.2 and 5.3 represents the outcome of a policy scenario, so the degree of fragmentations would be expected to reflect the effect of policy implication on the original landscape.

CHAPTER 4: DATA ANALYSIS RESULTS

In this chapter, brief analysis results based on the collected data are presented. While the analysis of the demographic data from the survey described the background of households with all associated information, the analysis of spatial data revealed the dynamics of land uses in the landscape. They all helped in better understanding the circumstances of land use decision making.

Tracing the history of land use changes from temporal satellite images helped to create the land use trajectories of the landscape. Moreover, the spatial data were associated with the land use trajectories to better understand the land use changes over time and space. The power of spatial analysis was exposed in overlaying and ranking the satellite images to define a new characteristic of the landscape: the resistance to land use change. Any location of interest in the landscape which had not been changed over period 1995-2010 could be considered stable from being changed.

4.1. Analysing demographic data

The surveyed data of 181 households in the study area were inputted into a Microsoft Access database for further analysis. Based off those data, some facts of demography can be drawn to give an overview about the My Lam commune. Most households in this survey are migrants. 178 or 98.3% of them migrated from the northern provinces of Vietnam during the 1990s following the New Economic Zone programme. This programme was designed by central government with the intention to populate and develop the Central Highlands of Vietnam after the Vietnam War. All interviewees could read or write because they finished at least secondary school before migrating to this area. Details of education level are presented in Figure 19.

Figure 19. Education level of interviewees

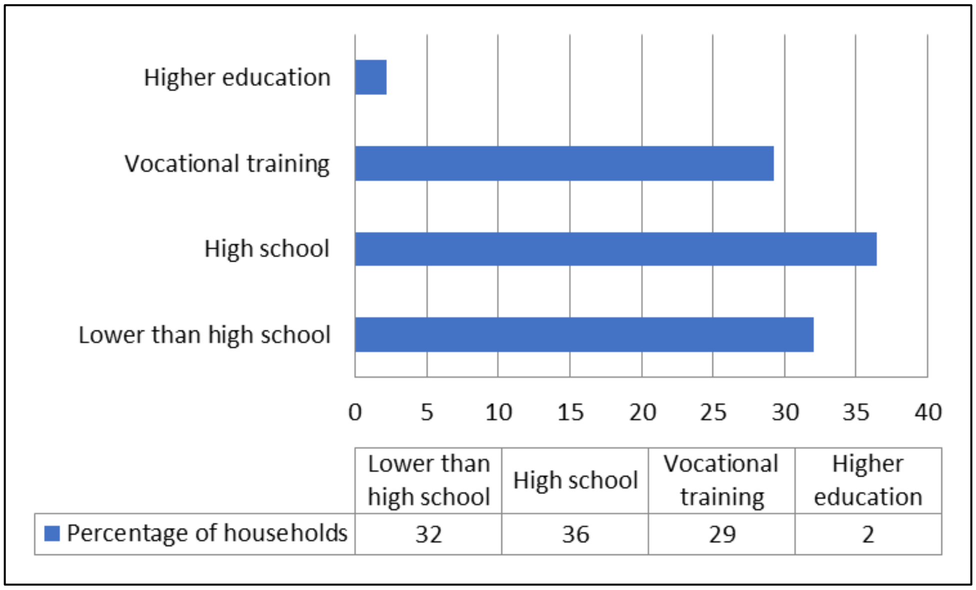
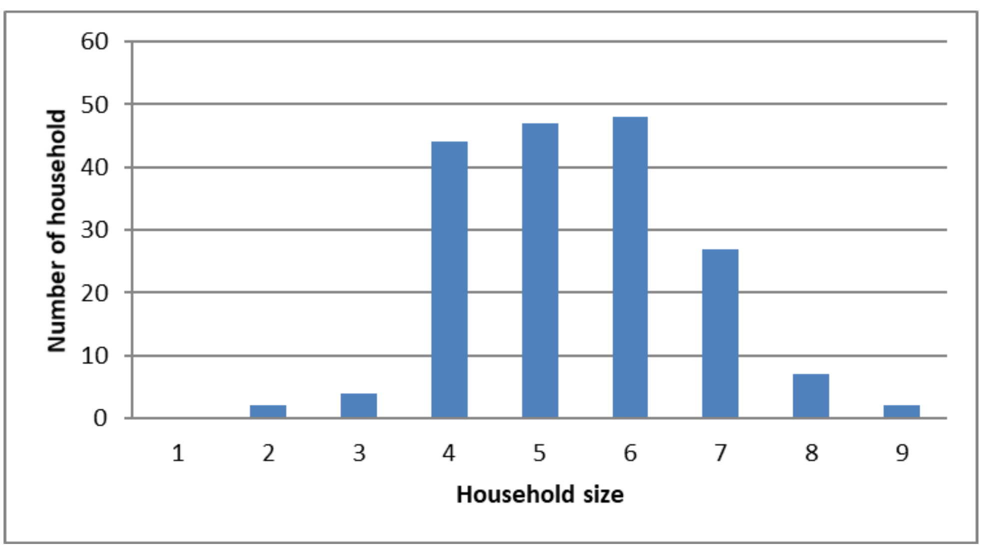


Figure 19 shows that 36% of households finished high school level (17 years of age) and 32% of them went further for vocational training. People have to stay in this area and work in agriculture for many reasons, such as the limitations of the labour market, the distance to centres, family matters, etc.

Responses from the 181 households show that the average household size (or the number of people living in one household) is 5.4 persons, with a maximum of 9 persons and a minimum of 2 persons. The histogram in Figure 20 gives more details of household size distribution among the sample size of 181.

Figure 20. Histogram of household size



According to the data in Figure 20, household sizes of 4, 5 and 6 persons are quite common in this commune. They share approximately 77% of total sample size. The second most common is a 7-person household size with 16.6% of the sample size. Small sized households of 2 and 3 persons share a small part of the total 181 households. Oversize households with 8 or 9 persons under the same roof are much fewer compared to the number of 4 to 6 person households. According to the General Statistics Office of Vietnam, the average household size for Central Highlands in 2009 was 4.8 persons; less than average of 5.4 persons found for this survey. However, the results show a general trend that household size in rural areas in Vietnam is rather high compared to the national level of 3.8 persons in 2009 (Binh 2011).

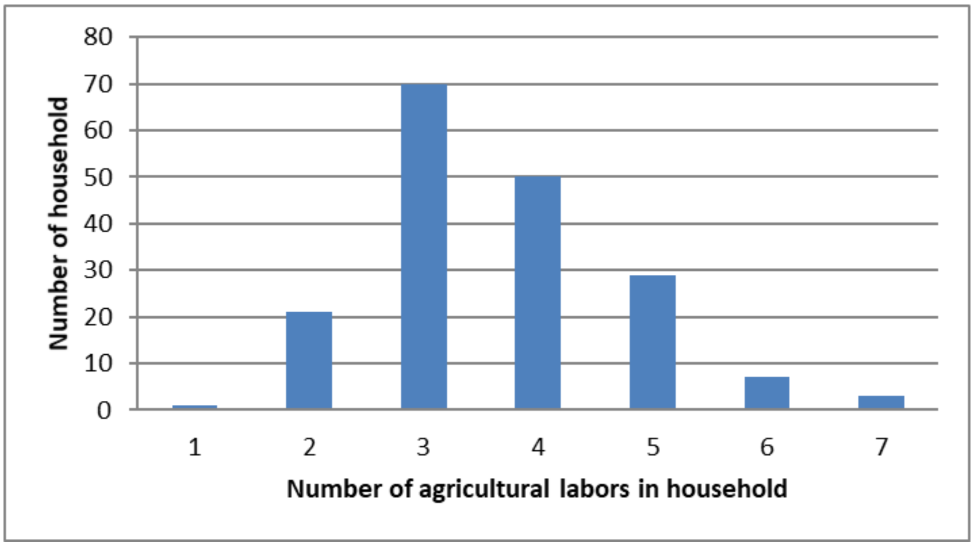
In the study area, it is quite common for many generations to live together, which helps to explain the relatively big household sizes. Senior people in each household are considered as heads of the household and they may have their adult children living together. One of the demographic questions in the questionnaire was designed to capture the age of the oldest child in each family, with the intention of identifying the age of the labour force in each household. The age distributions are summarized in Figure 21 below.

Figure 21. Distribution of age of the eldest child in households

Figure 21 shows that children in the study area start helping their families quite early with a few agricultural activities even before they reach age 15. Elder children from 16 to 22 may continue to study up to high school or vocational training then leave to work or pursue higher education. Young people aged from 23 to 45 in this survey seemed to settle down in their local area and become involved in agriculture. A few of them must find jobs in other regions during the down time between crops but return home to support families during the harvesting seasons. In some multi-generational households, adults aged 45 or more still live with and support their parents. Results also show that family members aged from 16 to 45 years old, who could actively be involved in agriculture, share a major 66% of the total sample size.

For more details about the labour force in each household, respondents were asked to estimate the number of full time equivalent (FTE) agricultural labours. Teenagers who have already been partially involved in agriculture were taken into account. For example, some households considered their 15 year old children as 0.5 FTE labour based on the hours or days they could support their family. The histogram of agricultural labour force represented in Figure 22 reflects the responses from households.

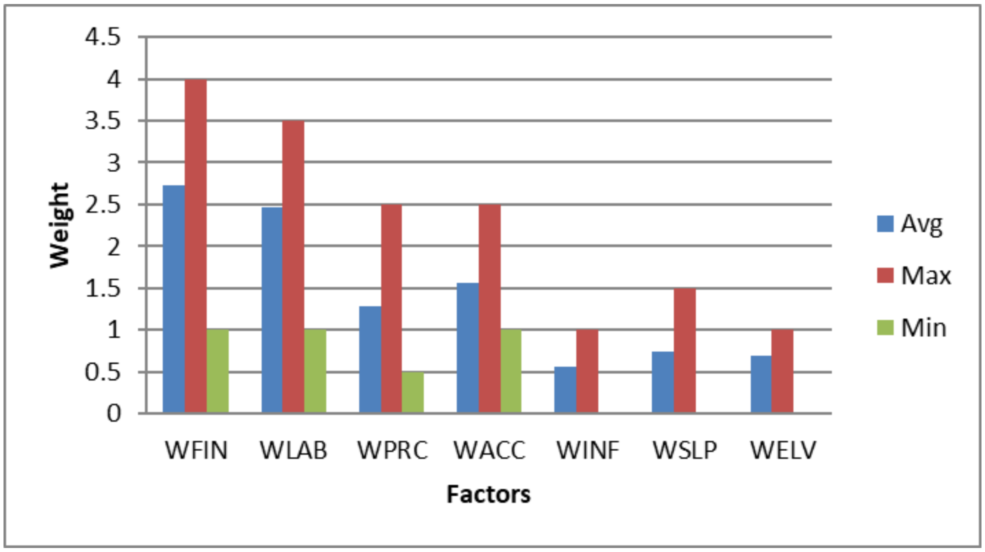
Figure 22. Histogram of agricultural labours in households



As presented in Figure 22, households with at least 3 FTE are the majority. Groups of 4 and 5 FTE households are also popular among the samples. About 10% of surveyed households have 2 FTE or less. A few of the larger households had 6 or 7 FTE.

The collected data also showed an interesting fact of how households considered different demographic and geographical factors in their land use related decision making. This question listed 8 factors for households to weight. Factors used in the questionnaire were "Financial factor - WFIN", "Availability of labour in household - WLAB", "Price of commodities - WPRC", "Accessibility to infrastructure -WACC", "Influence from neighbours - WINF", "Influence from slope - WSLP", "Influence from elevation - WELV". Households had to determine how they weight those factors by recalling the way they made land use decisions. The grand total of all weights had to be 10. The summary is represented in Figure 23 below.

Figure 23. Weights for different factors in decision making



In Figure 23, it is clear that two demographic factors (WFIN and WLAB) have high average weights of 2.72 and 2.46 respectively. They also have high maximum weights compared to other factors. WFIN was weighted to 4 and WLAB was weighted to 3.5 by a few households.

4.2. The resistance to change value and probability of being unchanged (Z value)

For some locations in the study area, land use changed gradually but not that often at others. It is necessary to address why land uses were stable at some specific locations. Resistance to change was observed over the years from 1995 to 2010 using the satellite images. This resistance is difficult to explain because it is the result of a complex set of factors, but it can be examined with the raster algebra application in GIS.

To quantify the resistance to land use change throughout the years from 1995 to 2010, a measurement was designed based on the four processed satellite images of 1995, 2000, 2005 and 2010. The original images were obtained from SPOT 2 and SPOT 5 and processed by Space Technology Institute, Vietnam Academy of Science and Technology. Those processed images were divided into three 5-year sets: 1995 to 2000, 2000 to 2005 and 2005 to 2010. ArcGIS 10.1 was deployed to calculate the difference in value (indicating land cover) of each pixel from the beginning to the end of each period. The outputs of those calculations are binary surfaces. For instance, in the 1995 to 2000 period, at each pixel of interest, if land use from the year 2000 is different from land use of the year 1995 then the value of surface 95_00 at that pixel will be "1"; otherwise it is "0". After deriving 3 binary surfaces of change for the three periods, the nested algebra equation below was used to produce a raster of resistance to the land use changes:

[Resistance raster] = con([95_00] == 0 & [00_05] == 0 & [05_10] == 0, 7,

con([95_00] <> 0 & [00_05] == 0 & [05_10] == 0, 6,

con([95_00] <> 0 & [00_05] <> 0 & [05_10] == 0, 5,

con([95_00] == 0 & [00_05] == 0 & [05_10] <> 0, 4,

con([95_00] <> 0 & [00_05] <> 0 & [05_10] == 0, 3,

con([95_00] <> 0 & [00_05] == 0 & [05_10] <> 0, 2, 1))))))

This algebra equation expressed the algorithm used to evaluate the resistance to change at each pixel. Each conditional statement (con) compares values of a pixel taken from 95_00, 00_05 and 05_10 surfaces and all possible situations were stacked in a hierarchical ranking system from high to low values. The output of this statement was a raster named Resistance raster which reflects the stability of change at each pixel. The pixel size of this combined raster is 20m x 20m, the same as the original satellite images.

The rule for ranking the resistance to change is defined as pixels with no changes through all three periods (or a 0 value in each of the three periods) will have the highest stability or resistance and carry a value of 7, while pixels that have changed over each of the three periods (or a 1 value in each of the three periods) it is considered as “easy to change” and gets value of 1 - the lowest resistance to change. Other intermediate cases get values from 2 to 6 on this scale. If a pixel did not change in the first period but then actively changed in the last two periods then it could carry value 2, quite a low resistance to change. If a pixel only changed once in the first period but then did not change in the last two periods then it could get value 6, a rather high stability level. The principle is that the more stable the pixel is the less likely its land use will change. It should be noted that this kind of ranking is calculated based on spatial data from 1995 to 2010 and in this research, it is assumed to represent the long-term state of the whole study area. All cases and possibilities are represented in Table 7 below.

Table 7. Rank of Resistance to change and Probability of being unchanged (Z value) at each pixel

Period	Cases						
1995-2000	0	0	0	1	1	1	1
2000-2005	0	0	1	1	0	1	0
2005-2010	0	1	1	1	0	0	1
Rank of resistance	**7**	**4**	**2**	**1**	**6**	**5**	**3**
Probability of being unchanged (Z)	**1.00** =1-(1/3)*0	**0.67** =1-(1/3)*1	**0.33** =1-(1/3)*2	**0** =1-(1/3)*3	**0.67** =1-(1/3)*1	**0.33** =1-(1/3)*2	**0.33** =1-(1/3)*2

The calculation in Table 7 shows that there are 2^3-1 = 7 possibilities or cases for a pixel to change its status from 1995 to 2010.

The probability of being unchanged (its Z value) was calculated from the rank of resistance using the following equation:

Equation 3:

$$Z = 1 - (1/\mathbf{p})*\mathbf{c} \qquad \text{with } Z = (0,1)$$

where **p** is the number of periods in ranking and **c** is the number of periods having land use changes (or carrying "1" in binary surface).

If there were no changes during the three periods (rank of resistance equals 7), **p** = 3 in this case is equal to 3 (or three periods) while **c**= 0 because there was no change in any period, and Z = 1- (1/3)*0 = 1. This case has the highest probability of being unchanged. If rank of resistance is equal to 1 or there were continuous changes during the three periods, Z = 1 - (1/3)*3 = 0 or the probability of being unchanged is lowest and this means land use is more likely to change.

The design of Z values as represented in Table 7 is to show where the land use covers have frequently changed and where they have not. The further implementation of those values is to quantify how easily a land cover at a specific location may change during the simulation process of the SeABM. The decision-making core of this agent-based model hypothesized that by comparing the resistance value to an index or figure, which is generated by the econometric model, a household could more easily make a land use decision at a specific location (or pixel in this research).

4.3. Analysis of land use trajectories

Identifying and analysing the LULC trajectories at the pixel level required getting the data of land use cover in each satellite image at the pixel level to compare and build up the LULC trajectories. It was also desirable to link the trajectories to other spatial variables, such as distance to river and road networks, elevation and slope at points of interest, to understand more about the underlying factors of LULC.

The feature classes of river and road networks were firstly converted to binary raster surfaces, from which Euclidean distance surfaces were derived using the built-in distance tools of ArcGIS.

The spatially-explicit representations of changes at the pixel level were collected and analysed through a set of sample points. There were 1000 random points designed and equally distributed to different strata or land use categories of a classified land use image taken in 1995. The year 1995 was considered as the starting point where the land use land cover types were diversified, due to low anthropogenic impacts from the early stage of resettlement.

The classified image from 1995 which covers My Lam commune was converted to polygon using "Raster to Polygon tool" to produce several single-part polygons of land uses. Those single-part polygons with the same land use codes located at different locations were then reunited into multi-part polygons. Each multi-part polygon represents a stratum of land cover. These 10 strata are shown in Table 8.

Table 8. 1000 equally random sample points divided by strata, based on land use land cover of 1995

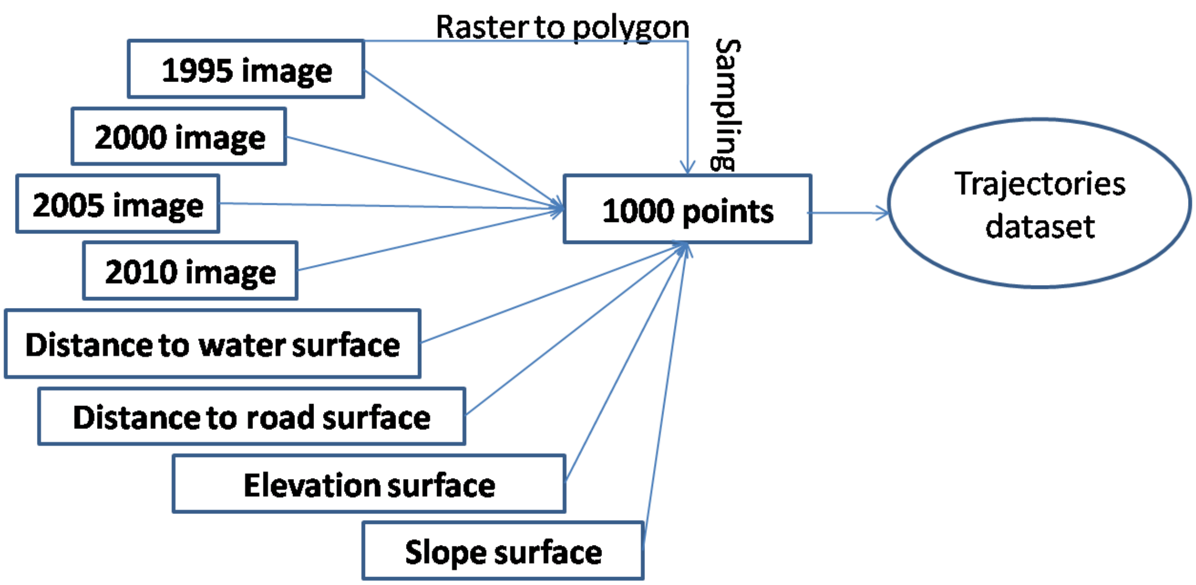
Stratum	LU code	Land use	Projected number of samples	Actual number of samples
1	2	Medium evergreen forest	100	89
2	3	Poor evergreen forest	100	96
3	5	Mixed forest (broad leaf with conifer or bamboo)	100	98
4	6	Bamboo forest	100	124
5	9	Rice	100	107
6	10	*Perennial trees (cashew - LU code 101, acacia hybrid - LU code 102)	100	96
7	11	*Annual crops (maize - LU code 111, cassava - LU code 112)	100	104
8	16	Settlement	100	92
9	17	Water body	100	98
10	18	Young forest (re-growth forest)	100	96

*LU - land use*LU code for Perennial trees and Annual crops were subdivided and recoded*

As represented in Table 8, for each stratum, 100 sample points were randomly generated with ArcGIS to create a point shapefile. The actual number of samples generated by ArcGIS was a slightly different from the designed one. This fact did not affect any further steps. Those 1000 points were used to extract the land use data, proximities to river and road networks, and elevation and slope from relevant surfaces from satellite images taken in other years. Although data from satellite images did not differentiate between cashew and acacia hybrid or between cassava and maize, the observation and local consultation from fieldtrips helped in identifying and recoding them, as presented in Table 8.

The procedure of data preparation and extraction for LULC trajectories is illustrated in Figure 24.

Figure 24. Sampling and data extraction process



As seen in Figure 24, sample points were used to extract values from different surfaces. Extracted data from satellite images were land use data which formed the raw LULC trajectories dataset. This dataset was then loaded into MS Excel to be analysed. The structure of the data analysis is represented in Table 9. This is a fraction of the analysis, which shows part of how LULC changed throughout the years.

Table 9. LULC change analysis based on 1000 sample points - a fraction of the full version

1995			2000			2005			2010		
LU	N	%	LU	N	%	LU	N	%	LU	N	%
2	89	100	2	47	53	2	26	29	3	10	11
									5	8	9
									6	8	9
						3	6	6.3	3	6	6.7
						5	7	7.3	6	7	7.9
						6	3	3.1			
						17	5	5.2	17	5	5.6
			3	10	11	3	3	3.1			
						5	4	4.2			
						17	3	3.1			
			5	27	30	3	4	4.2			
						5	21	22	5	1	1.1
									6	20	22
						6	1	1			
						17	1	1			
			6	5	5.6	3	1	1			
						6	3	3.1			
						17	1	1			
3	96	100	3	95	99	3	38	40	3	34	35
									5	4	4.2
						5	12	13	5	1	1
									6	11	11
						6	13	14	6	9	9.4
									10	1	1
									11	3	3.1
						10	1	1			
						17	5	5.2	17	5	5.2
						18	26	27	18	26	27
			5	1	1						

LU: land use code (Refer to Table 8) N: number of sample points %: percentage

In Table 9, starting from the 1995 data, the total number of sample points in each land use type (stratum) may remain unchanged or change to one or many land use types in the following periods (2000, 2005, 2010). The percentages of changed points are calculated over the years. From the data, there are some clear trends of how land uses could change in the following period. For example, there were some patches of "Medium evergreen forest" (LU = 2) in My Lam commune in 1995, represented by 89 sample points. By 2000, only 53% of points were kept as "Medium evergreen forest"; the other 47%

were converted to "Poor forest" (LU = 3), “Mixed forest” (LU = 5) or “Bamboo forest” (LU = 6). The “Poor forest” portion of the landscape seems to be stable through 15 years, that is why at least 35% of sample points for "Poor forest" remain in 2010. There were no sample points of "Medium evergreen forest" that remained until 2010. They were either converted to other land use types or downgraded to "Poor forest".

The data analysis was not only for capturing the land use changes in each year of interest, but also summarising the spatial conditions related to each land use change in each five-year period. The spatial variables associated with the land use changes are shown in Table 10.

Table 10. Land use change with associated spatial factors - a fraction of the full version

1995			2000			Elevation (m)		Slope (degree)		Distance to water (m)		Distance to road (m)	
LU	N	%	LU	N	%	max	min	max	min	max	min	max	min
2	89	100	2	47	53	242	23	70	6	293	0	3252	948
			3	10	11	182	35	73	19	50	0	3231	1709
			5	27	30	177	122	68	10	40	0	1821	1648
			6	5	5.6	160	44	69	47	20	10	3208	2769
3	96	100	3	95	99	313	19	73	3	427	0	3297	10
			5	1	1	145	145	21	21	189	189	402	402

LU: land use code (Refer to Table 8)   N: number of sample points   %: percentage   m: metre(s)

Table 10 shows that the sample points representing “Medium evergreen forest” (LU = 2) could remain constant during 1995 to 2000 if they carried certain spatial conditions regarding the elevation, slope and the distances to water and transportation. Locations where the elevation was up to 242m, slopes steeper than 70%, at a distance of 293m from water sources, and in range of 948m to 3252m from the road network seemed to maintain their land cover type of “Medium evergreen forest”. However, this forest cover type could be changed to “Poor forest" (LU = 3) or “Mixed forest” (LU = 5) or “Bamboo forest” (LU = 6) if it met certain spatial conditions, especially having a low proximity to water sources. The periods of 2000 to 2005 and 2005 to 2010, which were not included in this fraction, used the same principle to explain the land use cover change from one year to the other year with the associated spatial data to characterize the changes over the years. This helped the model determine in which specific conditions a land use cover could be changed and what would likely be the new land use cover.

The data in Table 9 and Table 10 were extracted from the raw trajectories dataset using pivot tables in MS Excel 2007. For each five-year period, land uses of the starting and ending year were used as inputs to query values of several spatial variables and those values were sorted to get the maximum and minimum.

The results of data analysis based on 1000 sample points equal in number and randomly distributed spatially in 10 strata are summarised in Table 11 below. This is the representation of LULC trajectories in the study area from 1995 to 2010 with at least 5% of change.

Table 11. Land use cover change trajectories from 1995 to 2010

1995		2000		2005		2010	
Land use cover	**%**	**Land use cover**	**%**	**Land use cover**	**%**	**Land use cover**	**%**
Medium forest	100	Medium forest	52.8	Medium forest	29.2	Poor forest	11.2
						Mixed forest	9.0
						Bamboo forest	9.0
				Poor forest	6.3	Poor forest	6.7
				Mixed forest	7.3	Bamboo forest	7.9
				Waterbody	5.2	Waterbody	5.6
		Poor forest	11.2				
		Mixed forest	30.3	Mixed forest	21.9	Bamboo forest	22.5
		Bamboo forest	5.6				
Poor forest	100	Poor forest	99.0	Poor forest	39.6	Poor forest	35.4
				Mix forest	12.5	Bamboo forest	11.5
				Bamboo forest	13.5	Bamboo forest	9.4
				Waterbody	5.2	Waterbody	5.2
				Young forest	27.1	Young forest	27.1
Mixed forest	100	Poor forest	10.2	Poor forest	5.1	Poor forest	5.1
		Mixed forest	75.5	Mixed forest	50.0	Bamboo forest	50.0
				Bamboo forest	19.4	Bamboo forest	7.1
						Other perennial trees	8.2
		Bamboo forest	10.2	Bamboo forest	10.2	Annual crops	7.1
Bamboo forest	100	Bamboo forest	87.9	Bamboo forest	70.2	Bamboo forest	35.5
						Other perennial trees	18.5
						Annual crops	10.5
				Other perennial trees	5.6	Other perennial trees	5.6

1995		2000		2005		2010	
Land use cover	**%**	**Land use cover**	**%**	**Land use cover**	**%**	**Land use cover**	**%**
Rice field	100	Rice field	64.5	Rice field	61.7	Rice field	60.7
		Other perennial trees	11.2	Other perennial trees	8.4	Other perennial trees	8.4
		Annual crops	6.5	Annual crops	5.6		
		Bare land	15.0	Bare land	14.0	Bare land	13.1
Other perennial trees	100	Rice field	11.5	Annual crops	5.2		
		Other perennial trees	43.8	Bamboo forest	22.9	Other perennial trees	17.7
						Annual crops	5.2
				Other perennial trees	15.6	Other perennial trees	15.6
				Annual crops	5.2		
		Annual crops	44.8	Bamboo forest	7.3		
						Annual crops	6.3
				Annual crops	33.3	Annual crops	32.3
Annual crops	100	Bamboo forest	35.6	Bamboo forest	21.2	Rice field	8.7
						Other perennial trees	10.6
				Other perennial trees	5.8	Other perennial trees	5.8
		Rice field	32.7	Rice field	27.9	Rice field	27.9
		Other perennial trees	11.5	Other perennial trees	9.6	Other perennial trees	9.6
		Annual crops	6.7				
		Bare land	10.6	Bare land	7.7		
Built-up area	100	Rice field	52.2	Rice field	51.1	Rice field	51.1
		Other perennial trees	22.8	Other perennial trees	19.6	Other perennial trees	19.6
		Annual crops	6.5	Annual crops	6.5	Annual crops	6.5
		Built-up area	15.2	Built-up area	10.9	Built-up area	10.9

1995		2000		2005		2010	
Land use cover	**%**	**Land use cover**	**%**	**Land use cover**	**%**	**Land use cover**	**%**
Waterbody	100						
		Rice field	16.3	Rice field	14.3	Rice field	14.3
		Bare land	25.5	Bare land	20.4	Bare land	20.4
		Waterbody	46.9	Bamboo forest	7.1	Bamboo forest	5.1
				Waterbody	29.6	Waterbody	29.6

The LULC trajectories start from the left to the right of Table 11 following the time series of satellite images available for this study. At every 5-year period, all the changes in land use cover at sample points were aggregated into groups with percentage of each change, which can be considered as the intensity of change. One land use cover type from 1995 can be changed or converted to different types depending on its initial type under either natural successions or anthropogenic activities. In other words, Table 11 shows how each land cover type evolves over time.

In general, good forest cover in the study area has a downgrading trend. Only about 53% of "Medium evergreen forest" points were kept from year 1995 to 2000 while 47% of those points were gradually converted to "Poor forest", "Mixed forest" and "Bamboo forest". Only 29% of "Medium evergreen forest" points were left in 2005 and all of them were degraded to poor, mixed and bamboo forest in 2010.

The "Poor forest" type also has a downgrading trend while more than 35% of "Poor forest" cover points were kept until 2010, 65% of them were converted or degraded to other lower quality forest types. A significant number of points originating as "Bamboo forest" were converted to other land uses, mainly to perennial and annual crops. Only 35.5% of points maintained their status type of "Bamboo forest" in 2010.

About 60% of points having rice land status were kept over the 15 years. This shows that paddy rice cultivation still plays an important part in agricultural activities. There was a trend to change suitable "Rice field" to annual and perennial crops, which reflected the diversification of cultivation in later years after settlement. For some reason, a portion of rice field has been abandoned as "Bare land". This could be explained by the proximity, the reduction in soil quality and the changes in agricultural practice. There is a clear interchange among land use types such as "Rice field", "Perennial crops" and "Annual crops" through the years.

It is difficult to identify the reasons or mechanisms behind LULC changes due to the complexity of local land use practices and socioeconomic development. Data from satellite images themselves only shows snapshots of the landscape from which interpretations regarding the changes and their trends can be made. However, the appearance of land use changes at specific locations may be driven by some spatially related factors. By incorporating those spatial data to the LULC trajectories, the picture of landscape dynamic becomes more informative. Table 12 describes the relationship between location's geographic features and LULC trajectories from 2000 to 2010.

Table 12. LULC trajectories from 2000 to 2010 and associated spatial related features

From 2000 to 2005	Elevation (m)		Slope (percent rise)		Distance to hydrology (m)		Distance to road network (m)	
	max	min	max	min	max	min	max	min
Medium forest to Poor forest	242	170	51	21	251	0	2843	994
Medium forest to Mixed forest	217	151	50	6	293	10	2992	948
Medium forest to Waterbody	44	23	70	35	10	0	3247	3153
Poor forest to Mixed forest	267	28	73	24	344	10	2004	136
Poor forest to Bamboo forest	248	148	41	3	338	64	2191	30
Poor forest to Young forest	313	144	38	9	334	30	3297	418
Mixed forest to Bamboo forest	261	141	39	10	423	0	2973	5
Bamboo forest to Other perennial trees	191	128	34	2	322	36	555	14
Rice field to Annual crops	134	133	3	2	51	45	53	35
Other perennial tree to Bamboo forest	144	132	31	1	349	80	205	79
Other perennial tree to Annual crops	132	132	2	1	70	60	99	88
Annual crops to Bamboo forest	134	132	3	2	91	59	81	34
Waterbody to Bamboo forest	130	31	74	50	14	0	3234	2562

From 2005 to 2010	Elevation (m)		Slope (percent rise)		Distance to hydrology (m)		Distance to road network (m)	
	max	min	max	min	max	min	max	min
Medium forest to Poor forest	235	150	49	15	36	10	2822	2553
Medium forest to Mixed forest	166	23	64	12	20	10	3252	2978
Medium forest to Bamboo forest	242	210	41	27	290	198	1038	954
Mix forest to Bamboo forest	300	105	73	6	681	0	3118	15
Bamboo forest to Rice field	138	20	59	0	282	36	501	25
Bamboo forest to Other perennial trees	321	75	75	1	469	20	979	5
Bamboo forest to Annual crops	280	117	66	1	359	0	3320	54

m - meter(s)

As reported in Table 12, the 10-year period from 2000 to 2010 was subdivided into two 5-year sub-periods.

a) From 2000 to 2005:

"Medium forests" had been continuously converted to other poorer types of forest. Data show that from 2000 to 2005 "Medium forest" at locations varying from 171m to 242m, with moderate slopes up to 51%, and within 251m of the river network was converted to "Poor forest". It is worth noting that "Medium forest" close to the road network (proximity less than 1000m) seems to be the exception for the conversion to the "Poor forest". "Medium forest" with similar features but located at flatter slopes from 6% was converted to "Mixed forest". "Medium forest" with low elevation and close to rivers was changed to "Water body" due to floods. However, the last type of conversion was only found at a relevant distance from the road network.

In this period, it was observed that "Poor forest" was upgraded to "Mixed forest" when it was located not too far from the road network and had elevations above 28m and slopes from 24%. At other locations with elevations higher than 144m and lower slopes from 3%, "Poor forests" were either downgraded to "Young forest" or converted to "Bamboo forest".

"Bamboo forest" having preferable spatial features was converted to "Perennial crops". Paddy rice within a short distance to road and water and a flat slope was converted to "Annual crops". "Perennial crops" were either converted to "Bamboo forest" or "Annual crops" when they were on flat slopes. At some locations, "Annual crops" and "Water body" were converted to "Bamboo forest". The first conversion happened on flat slopes while the second one happened on a steeply slope with a significant distance to the road network.

b) From 2005 to 2010:

The pattern of LULC in this sub-period looks simpler than the previous sub-period. The first observation is that "Medium forests" were still converted to "Poor forest", "Mixed forest" and "Bamboo forest" but at a further distance of at least 954m from the road network. "Medium forest" close to water was downgraded to "Poor forest" and "Mixed forest", while locations further than 198m were converted to "Bamboo forest".

"Mixed forest" was converted to "Bamboo forest" with wide ranges of spatial features. This conversion was observed at locations with elevations from 105m to 300m, with slopes varying from 6% to 73%, and within a distance of 681m from water and 3118m from road networks.

"Bamboo forest" in this sub-period has three types of conversion. It was converted to "Rice field" if the location was lower than 138m elevation and within 282m of a source of water and 501m from the road network. At a further distance to water and road networks, with moderate elevation, "Bamboo forest" was converted to "Annual crops". Those locations with high elevations were converted to "Perennial crops".

Representations of Table 11 and Table 12 suggest that the LULC trajectories become simpler throughout the years. Those tables reveal that the diversity and complexity of LULC conversions reduced over the years; there were only seven types of typical conversion found in 2005-2010. It should be noted that only conversions of more than 5% from the original sample points were considered for the data analysis, so there could be some minor types of conversion that happened on the landscape that were not counted.

A massive conversion of different forest covers to other types of land use was found in the analysis during this period. It reflects an active period for agricultural extension right after the resettlement of migrants. In this active period, there were interchange conversions among crops, especially the perennial and annual crops as reactions to market demand and local food demand.

In section 5.1 the econometric model has used data of two periods, 2000-2005 and 2005-2010, to find the correlation between the LULC and other demographic and spatially related factors with an assumption that those periods have the latest data to support the research. Following this idea, LULC trajectories from 2000 to 2010 were considered as typical LULC trends for the projected period of 10 years starting from 2010.

4.4. Land use conversion modules

The question of what a LULC will change to can be answered by using LULC change modules to describe the process. Each module simulates the typical and popular land use change from one to other types based on the land use trajectories, and this mechanism was coded in simulation script.

The information taken from satellite images and consultation with local people showed that some LULC changes may take several years to occur while others may happen within a year. For example, it may take about five years for a "Medium forest" to be gradually degraded into a "Poor forest". Of course, a forest fire or timber harvesting could turn a forest into a clear site quickly, but those cases are not considered in this modelling. Conversions from "Bamboo forest" to production purposes can be done in a year. It is also quite easy to switch between rice and annual crops or between some perennial crops and annual crops.

The summary of data from Table 11 and Table 12 suggests a list of 20 possible LULC change modules. They are referred to as pre-set modules and numbered from 1 to 20 in Table 13 below.

Table 13. Land use land cover change modules

Land use cover changes	Module No.	Year(s) of fulfilment
Medium forest to Poor forest	1	5
Medium forest to Mixed forest	2	5
Medium forest to Water body	3	5
Poor forest to Mixed forest	4	5
Poor forest to Bamboo forest	5	5
Poor forest to Young forest	6	5
Mixed forest to Bamboo forest	7	5
Bamboo forest to Other perennial trees	8	1
Rice field to Annual crops	9	1
Other perennial trees to Bamboo forest	10	1
Other perennial trees to Annual crops	11	1
Annual crops to Bamboo forest	12	5
Water body to Bamboo forest	13	5
Medium forest to Poor forest	14	5
Medium forest to Mixed forest	15	5
Medium forest to Bamboo forest	16	5
Mixed forest to Bamboo forest	17	5
Bamboo forest to Rice field	18	1
Bamboo forest to Other perennial trees	19	1
Bamboo forest to Annual crops	20	1
NO CHANGES: age increased/forest type upgrade	0	1

Table 13 represents the full list of LULC change modules which were implemented during the simulation. Beside 20 pre-set LULCs, there is a module numbered "0" which describes the normal change of the LULC in every year (or business as usual). For example, areas with perennial trees and forest trees get older after each year if there are no changes to that area. This information was accumulated during the simulation and controlled for in an auto upgrading of the related LULC. For instance, after five years, a "Young forest" area will be upgraded to "Mixed forest" if no other changes have occurred.

It is worth noting that from 2000 to 2010 some LULC happened in the first five years (2000-2005) and others happened in last five years (2005-2010). According to this fact, modules from 1 to 7 and from 12 to 17 were activated only when the current counter of year of simulation was higher than 5 years.

4.5. Household profiles and land use decision making

The next important question to answer for simulation is which household will initiate which LULC change module(s)? Based on data from the survey, the households' demography seems to be heterogeneous and they may have diversified land use under their management. However, their household decision-making behaviour appeared to be determined by two main factors: cash availability (**cashrate**) and labour availability (**labrate**) in each household. Which agricultural practice a household may choose depends on how much it can invest after considering the living expenditure for the whole family and whether it has enough labour to carry out that practice. So, what are the threshold values for cash and labour availabilities for a household's decision making? Discussions with local people revealed that they made different decisions on land use change depending on whether **cashrate** and **labrate** were negative or positive. In other words, those rates are compared against 0. The formulas to calculate those rates can be written as below:

cashrate = **cashbal**/(**nonagrexp+ agrinv**)

where **cashrate** is the rate between the net cash balance of a household and its total expenditure (which includes non-agricultural expenditure and investment into agricultural activities). In case the cash balance of a household is less than 0 and it needs to borrow money, then:

cashrate = (**cashbal** + **borrow**)/(**nonagrexp** + **agrinv**)

where **borrow** is the amount of cash that household can borrow to add up to its cash availability. In this research, it was set at NZ$ 300 based on discussions of household average credits with informants. However, sometimes the cash availability is still less than 0 after adding up the borrowed money.

In any case, the **labrate** of a household is:

labrate = **labbal**/(**hhlab***20 days *12 months)

where **labrate** is the rate between the labour balance (**labbal**) of a household with its total labour force, which can be calculated from the number of labours in household (**hhlab**) multiplying by working days per month (20 days) and then multiplying by the 12 months of the year.

Since agricultural activities depend on the cultivation regime of each crop, the requirement of labour for each crop in each month can be different. Crops' information is provided in Table 14 below.

Table 14. Main crops and activities with their schedule during a year

Crops/Activities	Standard unit	Annual labour/unit	Months of activity											
			1	2	3	4	5	6	7	8	9	10	11	12
Paddy	1 ha	150	x	x	N	N	N	x	x	x	N	N	x	x
Cashew	1 ha	50	x	x	x	x	x	x	x	x	x	x	x	x
Acacia hybrid	1 ha	60	x	x	x	x	x	x	x	x	x	x	x	x
Aquaculture	0.25 ha	52	N	N	x	x	x	x	N	N	x	x	x	x
Cassava	1 ha	130	x	x	N	N	N	x	x	x	N	N	x	x
Maize	1ha	100	x	x	N	N	N	x	x	x	N	N	x	x
Forest protection	1 ha	10	x	x	x	x	x	x	x	x	x	x	x	x
Bamboo shoots collection	1 ha	10	x	x	x	x	x	x	x	x	x	x	x	x

x – labour is needed for activity in current month　　*N - no labour needed*

Table 14 lists eight basic crops and activities that occupy these areas and can be manipulated in the simulation. Activities such as husbandry are not area-based so they were excluded from the simulation. Table 14 also provided data on how busy those eight activities are in a calendar year. Few crops and activities require monthly labour, however households cultivating cassava have 5 months off. In this case, labour balance of a household can be calculated by following equation:

Equation 4:

$$\mathbf{labbal} = \sum \boldsymbol{labbal}(\omega)$$

where **labbal** is total labour balance of a household in a year and ***labbal***(ω) is the labour balance of each crop or activity listed in Table 14, in a year.

The ***labbal***(ω) can be obtained by deducting the total labours needed for each type of crops or activities from the total labour force of a household in each month of 12 months of a year as in the below equation.

Equation 5:

$$\boldsymbol{labbal}(\omega) = \sum_{i=1}^{12}(\boldsymbol{hhlab} * 20 - \boldsymbol{lab}(\omega))$$

where (***hhlab****20) is the total labour availability per month of a household, using the assumption that each labourer in the household can work 20 working days (8 hours/day) in a month. ***lab***(ω) is the labour needed for each crop or agricultural activity in a month calculated by dividing total labour required for crops or activities by number of months those crops are cultivated or carried out. The **cashrate** and **labrate** help in building up the matrix of households' profiles by comparing those values to the threshold value of 0. The allocation is represented in Table 15.

Table 15. Households' profiles based on their cash and labour force availability

cashrate\labrate	surplus	deficit
surplus	Profile 1: **cashrate** and **labrate** are higher than 0. Households in this profile have more than enough resource to change the land use, invest in expensive crops like cashew, or convert poor forest to more profitable land use.	Profile 2: **cashrate** is higher than 0 while **labrate** is equal to or less than 0. Household has cash but no labour availability to get involved in labour intensive activities. Household keeps low labour demand activities and can invest into higher profit crops.
deficit	Profile 3: **cashrate** is equal to or less than 0 but **labrate** is higher than 0. In this situation, households can carry out some low-cost labour intensive activities, such as converting poor or bamboo forest to other land use to get a quick return.	Profile 4: both **cashrate** and **labrate** are equal to or less than 0. Household has limitations in investment and some high cost crops such as perennial crops may be abandoned or converted to annual crops if they are low investment and low labour intensive.

In Table 15 there are 4 profiles based on the balance of cash and labour in each household. Those profiles were used to group households (or agents) into different subgroups, with the assumption the decision making of households in each subgroup is homogenous. Households in "Profile 1" have more than enough resources to carry out any high cost and labour-intensive crops or activities while those in "Profile 4" have some limitations and may have to switch from a long-term investment into short-term crops/activities with a low labour intensity to cope with their situation.

The profile designations and the list of 21 potential LULC change modules were developed based on consultations with local informants, and LULC change modules allocated to suitable profiles. This means that a household belonging to a certain profile may have a set of LULC changes to make. The details of LULC change modules and their corresponding profiles is presented in Table 16.

Table 16. LULC change modules and their corresponding profiles

cashrate\labrate	surplus	deficit
	Profile 1	**Profile 2**
surplus	(0) No land use cover changes (3) Medium forest to Water body (8) Bamboo forest to Other perennial trees (18) Bamboo forest to Rice field (19) Bamboo forest to Other perennial trees	(0) No land use cover changes (3) Medium forest to Water body (4) Poor forest to Mixed forest (5) Poor forest to Bamboo forest (6) Poor forest to Young forest (7) Mixed forest to Bamboo forest
	Profile 3	**Profile 4**
deficit	(0) No land use cover changes (3) Medium forest to Water body (9) Rice field to Annual crops (14) Medium forest to Poor forest (15) Medium forest to Mixed forest (16) Medium forest to Bamboo forest (17) Mixed forest to Bamboo forest (20) Bamboo forest to Annual crops	(0) No land use cover changes (1) Medium forest to Poor forest (2) Medium forest to Mixed forest (3) Medium forest to Water body (10) Other perennial trees to Bamboo forest (11) Other perennial trees to Annual crops (12) Annual crops to Bamboo forest (13) Water body to Bamboo forest

example: (3) - LULC change module

As seen in Table 16, module (0) - "no changes" presents in all profiles in case households cannot make any change; the current LULC will be kept and updated year by year until the simulation stops. In the simulation process, some calculations will be made to identify the profile of a household and if there is a preferable condition to change a land use, it will strictly follow the predesigned LULC change modules. Since there was no previous information of household decision-making behaviour in this area, it is necessary to mention again how a household makes a decision is based on its profile and what associated conversion modules it may follow. This assumption helped in simplifying the mechanism of decision making in simulation. The content of these profiles and LULC change modules was discussed with the local people to make sure they sounded reasonable.

Chapter 4 provides detailed data analyses for this research. The demographic data from the survey described the background of households in the study area as well as the context of how local people live and make decisions about their land. The analyses of spatial data supplied rich information for the land use change to answer questions of what the future land uses are and how do conversions happen. They also generated parameters for SeABM simulation.

CHAPTER 5: RESULTS

Chapter 5 reports the results of the research. It starts with the output of the econometric model, which answers the first research question of what the main driving factors for environmental degradation are. From the output of the econometric model, the Y* value or the probability of being changed, was calculated last parameter needed for the SeABM simulation. The land use maps of simulations with different scenarios were represented and changes in land uses used as a proxy to explain other environmental degradation such as carbon sequestration, soil erosion and forest fragmentation. Those are the answers to the second and the third research questions.

5.1. Influencing factors of environmental degradation

5.1.1. Output of probit regression

Probit regression was chosen as an econometric model in this case to deal with the dichotomous dependent variable. The interpretation of the probit regression shows a set of spatial and non-spatial factors influencing land use cover change. The outputs of the probit regression are also important for the SeABM which was discussed in section 3.1. They inform the rule set that determines when and where a land use can be changed. From this point the SeABM shows the pace of change in land use covers, including forest coverage. Forest coverage predicted by the model can be used to project potential environmental degradation such as loss of carbon, soil erosion and landscape fragmentation.

Stata 10 has been used to estimate the parameters of this econometric model and the outputs are represented in Table 17.

Table 17. The outputs of probit regression

```
Probit regression                                 Number of obs   =        181
                                                  LR chi2(3)      =      95.77
                                                  Prob > chi2     =     0.0000
Log likelihood = -64.083121                       Pseudo R2       =     0.4277
```

chg0510	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
chg0005	-1.463567	.3145262	-4.65	0.000	-2.080027	-.8471072
elv	.0119636	.0033892	3.53	0.000	.005321	.0186062
agrlab	.8775443	.1374017	6.39	0.000	.6082419	1.146847
_cons	-4.443067	.7406385	-6.00	0.000	-5.894692	-2.991442

Table 17 shows the estimation of all predictors' regression coefficients and other parameters of the fitted model.

The probit regression created several outputs which allowed analysis interpretation, following the guidelines documented by UCLA Consulting Group and Stata.com.

The fitted model has a log likelihood of -64.083121. The higher the log likelihood values (close to zero) are, the better the model is fitted. This log likelihood value is used in the Likelihood Ratio Chi-Square (LR chi2) test of to find out whether all coefficients of **chg0005**, **elv** and **agrlab** in the model are simultaneously zero.

LR chi2(3) indicates the value of the Likelihood Ratio Chi-Square test with a degree of freedom of 3 (or 3 predictors) and at least one of the predictors' regression coefficient is not equal to zero. By looking for the **p-value** associated with the **LR chi2** value of 95.77 with a degree of freedom of 3, the **p-value** is less than 0.00001, which means the result is significant as **p** < 0.05 (at the 95% level). In other words, this leads to a rejection of the null hypothesis that all of predictors' regression coefficients are simultaneously equal to zero.

In this probit regression, McFadden's pseudo R-squared (**Pseudo R2**) does not have an equivalent to the R^2 or the proportion of variance of the response variable explained by the predictors that is found in Ordinary Least Squares (OLS) regression, according to the guideline from the UCLA consulting group's manual (UCLA 2014). Because this statistic does not mean what R^2 means in OLS regression, it is not correct to conclude that the fitted model can explain about 42.8% of observed data. In Table 17 the **Std. Err.** column shows the standard errors of the individual regression coefficients. They are used to calculate the **z** test statistic and the confidence interval of the regression coefficient.

The test statistic **z** is the ratio of the **Coef.** to the **Std. Err.** of each respective predictor. The **z** value follows a standard normal distribution, which is used to test against a two-sided alternative hypothesis that the **Coef.** is not equal to zero.

The parameter **P>|z|** is the probability the **z** test statistic (or a more extreme test statistic) would be seen under the null hypothesis that a predictor’s regression coefficient is zero, given that the rest of the predictors are in the model. For a given alpha level of 0.05, **P>|z|** determines whether or not the null hypothesis can be rejected. If **P>|z|** is less than 0.05, then the null hypothesis can be rejected, and the parameter estimate is considered statistically significant at 0.05 (or at confidence level of 95%). **P>|z|** values of predictors that are less than 0.05 means the estimated coefficients of the predictors are statistically significant and the null hypothesis (Ho) is rejected. This means three proposed predictors **chg0005**, **elv** and **agrlab** can be used to explain the dependent variable **chg0510**.

chg0510 is the binary response variable predicted by the model.

chg0005 -The **z** test statistic for the predictor **chg0005** is (-1.463567/0.3145262 =) -4.65 with an associated p-value (0.000) less than an alpha level of 0.05. The null hypothesis can be rejected, and this leads to a conclusion that the regression coefficient for **chg0005** has been found to be statistically different from zero given **elv** and **agrlab** are in the model.

The **z** test statistic for the predictor **elv** is (0.0119636/0.0033892 =) 3.53 with an associated p-value much less than an alpha level of 0.05. The null hypothesis would be rejected, and it can be concluded that the regression coefficient for **elv** has been found to be statistically different from zero given **chg0005** and **agrlab** are in the model.

The **z** test statistic for the intercept, **_cons,**is (-4.443067/0.7460358 =) -6.00 with an associated p-value much less than 0.05. The null hypothesis would be rejected, and it can be concluded that **_cons** has been found to be statistically different from zero, given **chg0005**, **elv** and **agrlab** are in the model and evaluated at zero.

Column **[95% Conf. Interval]** is the Confidence Interval (CI) for an individual coefficient given that the other predictors are in the model. For a given predictor with a level of 95% confidence, it could be stated that the "true" coefficient lies between the lower and upper limit of the interval with 95% confidence. It is calculated as ($\mathbf{z_{\alpha/2}}$)*(**Std.Err.**), where $\mathbf{z_{\alpha/2}}$is a critical value on the standard normal distribution or the **Confidence coefficient**. **α** is the confidence level, in this case 95% (or 0.95 in decimal format). The CI is equivalent to the **z** test statistic: if the CI includes zero, it would be false to reject the null hypothesis that a regression coefficient is zero, given the other predictors that are in the model. An advantage of a CI is that it is illustrative; it provides a range where the "true" parameter may lie within.

The estimated predictors seem to be technically fitted to the model. However, the signs of their coefficients may play a more important role in determining whether the predictors have logical meanings to the model.

elv has positive signs which intuitively appears to not be logical as elevation naturally should function as a negative factor for any man-made land use change, especially for agricultural purposes because the choice of crops depends on the elevation. However, in this case the positive sign of elevation shows that the higher the location and the steeper the slope, the higher the chance of land use cover being changed between 2005 and 2010. In fact, the topographic data and the ground truth information indicate that the maximum elevation of My Lam area is around 400m, which is suitable for many kinds of crops. And on the peaks, the red basaltic soil is much better than the clay soil on flat breaks. The positive sign turns out to be very logical if the explanation considers that the local farmers who migrated in the late 1980s had already converted nearly all the low land area to paddy rice to satisfy the food demand (1985-1990) and now they must go to higher elevations and steeper slopes to change the land use if needed. This means that **elv** can have a positive sign. Compared to the other two independent variables, **elv** has a weak impact on the dependent variable because the coefficient of 0.011 is far less than the other coefficients.

Other variables have their signs as anticipated. The variable **chg0005** with a negative sign shows that if land use cover at the current sample point has already changed in 2000-2005 then it will reduce the chance to be changed again in 2005-2010. The higher the number of the agriculture-involved labourers (**agrlab**) the higher the chance of a change in the land use cover in 2005-2010.

Except for the coefficient of the intercept (-4.443067), variable **chg0005** has a very strong influence on the dependent variable, with an absolute coefficient of 1.463567. **agrlab** also shows a significant level of influence on **chg0510**.

In conclusion, for the interpretation of the probit regression outputs, land use is less likely changed in the 2005-2010 period if it has been changed during the 2000-2005 period, but it is more likely to have changed if the elevation and the availability of labourers are high.

Based on the outputs of the probit regression, the probability of land use change in the 2005-2010 period can be obtained using Equation 6 as below:

Equation 6:

Y = Pr(**chg0510**) = -4.443067 - 1.463567***chg0005** + 0.0119636***elv** + 0.8775443***agrlab**

Where Pr(**chg0510**) is the probability of land use changing in 2005-2010, having value of 0 or 1.

Equation 6 represents a non-linear model which can be converted to standard linear model using a link function as presented in section 3.1.

5.1.2. Probability of land use change (Y* value)

The predicted probability of land use cover change in 2005-2010 can be calculated using the predicted coefficients as listed in Table 17. For a given point on the landscape, the predicted probability of its land use cover change can be calculated using Equation 7 which is adapted from Equation 2 in section 3.3.2.

Equation 7:

Y* = F(-4.443067 - 1.463567***chg0005** + 0.0119636***elv** + 0.8775443***agrlab**)

Where F is the cumulative distribution function of the standard normal.

The interpretation of the coefficients in this case is not as straightforward as the interpretation of coefficients in linear regression or logit regression. The change in probability attributed by a one-unit rise in a given predictor is dependent not only on its starting value but also on the values of the other predictors. For example, if **chg0005** and **elv** are held constant at zero, the one unit increased in **agrlab** from 2 to 3 has a different effect than the one unit increase from 3 to 4 as shown below:

With **agrlab** = 2: Y* = F(-4.443067 + 0.8775443***2**) = 0.0035943

With **agrlab** = 3: Y* = F(-4.443067 + 0.8775443***3**) = 0.03511425

With **agrlab** = 4: Y* = F(-4.443067 + 0.8775443***4**) = 0.17543843

The accumulated probability of land use cover change increases ten times (from 0.0035943 to 0.03511425) and five times (from 0.03511425 to 0.17543843) respectively.

The effects of a unit increase in **agrlab** are also strong if **chg0005** and **elv** are held at constant at their respective means (0.7679558 for **chg0005** and 147.0009 for **elv**) instead of zero. An example is presented as below:

With **agrlab** = 2:

Y* = F(-4.443067 - 1.463567*0.7679558 + 0.0119636*147.0009 + 0.8775443***2**) = 0.02002304

With **agrlab** = 3:

Y* = F(-4.443067 - 1.463567*0.7679558 + 0.0119636*147.0009 + 0.8775443***3**) = 0.11985161

With **agrlab** = 4:

Y* = F(-4.443067 - 1.463567*0.7679558 + 0.0119636*147.0009 + 0.8775443***4**) = 0.38278114

In this case, a one-unit increase in **agrlab** from 2 to 3 only results in 5.5 times change in predicted probability (from 0.02002304 to 0.11985161) and a one-unit change in **agrlab** from 3 to 4 only increases that predicted probability by three times. However, it is difficult to interpret these regression coefficients individually. The predicted probability increases when the predictor carrying a positive coefficient increases. But the increase of a negative coefficient may reduce the predicted probability. In this case, a rise in **elv** or **agrlab** is likely to increase the predicted probability of land use cover change in 2005-2010. **chg0005** or **_cons** of estimated model, in other words, decreases the predicted probability.

The coefficient of **_cons** is -4.443067. This means that if all predictors (**chg0005**, **elv** and **agrlab)** are zero simultaneously, the predicted probability of change in 2005-2010 at the location of interest will be:

$Y^* = F(-4.443067) = 0.000004434$

So, as expected, it is an extremely low predicted probability of land use cover change in 2005-2010.

The Y* value shows the predicted probability of land use cover change at a given location, and because it is a cumulative distribution function its value ranges from 0 to 1. In other words, the Y* value indicates the likelihood of being changed for a land use cover type, at a specific location, under the influence of different demographic and spatial factors. Those factors contribute to the land use decision-making process at the household level, which can result in land use cover changes at the landscape level.

The application of the predicted probability of land use change (Y*) in this research is important. Equation 7 shows that Y* values are dynamic because they depend on the factors in the right-hand side of the equation, some of which change over the time. Those values can be used to evaluate the possibility of a land use cover change by comparing with the probability of being unchanged (or stability to change) at the same location. In section 4.2, the probability of being unchanged during 30 years at the same location, named as Z value, also has values ranging from 0 to 1, and Z value is assumed as a constant at each location in the landscape. It can serve as a constraint to any land use cover change. This research hypothesizes that at any location of interest, if its Y* value is equal or greater than its Z value then the land use cover at that location has a high probability of being converted to a different type in the future. This is the key response to the question of when and where a land use cover could be changed in the simulation.

For the first research question, the econometric approach has proved to be an appropriate answer. Its outputs show that there is a combination of spatial and demographic variables influencing the predicted probability of land use cover changes. This predicted probability was assumed to have correlations to the change of land use in general, and forest cover in particular. Decreases in forest cover and changes from rich to poor forest reflected the present state of deforestation and forest degradation, which degraded the quality of the environment. The outputs of the econometric model pointed out that the past land use changes, the elevation and the agricultural labour force associated to a location of interest are the driving factors of land use changes, which serve as a proxy to evaluate potential environmental degradation.

5.2. Results for baseline scenario

In this research, it was a challenge to simulate the dynamics of any single environmental outcome due to the lack of representative data and the complexity of decision-making behaviour. Land use cover changes (forest cover change in this case) were used as a proxy to extrapolate environmental degradation; for example, when a rich forest is converted to a poor forest on a steeply slope, it is more susceptible to landslides and soil erosion. After having Y* value and other parameters calculated in different steps, they were fed into the Python scripts to start the simulation of future land use.

The Agent-based modelling (ABM) approach was deployed to simulate land use changes based on farmers' decision-making processes to answer the question of how some environmental outcomes could potentially be impacted in the future. Due to the integration of spatial data in this ABM, it is more precisely to be called a Spatially explicit Agent-based model (SeABM). It described when and how a household or agent makes its land use decisions against other factors throughout the projected period.

How a LULC change could be determined from different factors was described in section 5.1. Data used for the probit regression were from 2000-2010 so the simulation could be considered as the trend to forecast the LULC's dynamics in the next 10-year period from 2010 to 2020. Outputs of the econometric model played an important role in the simulation framework, as it indicated when agents would make a land use decision considering demography and spatial factors.

Since LULC's types are diverse, it is necessary to know what type of LULC will change and what it may change to. This question was answered using the analysis of the historical land use cover change patterns or land use cover change trajectories (section 4.3). By analysing the typical trajectories of land use changes in the study area, it was possible to embed them in the simulation framework to mimic the real situation happening in the landscape.

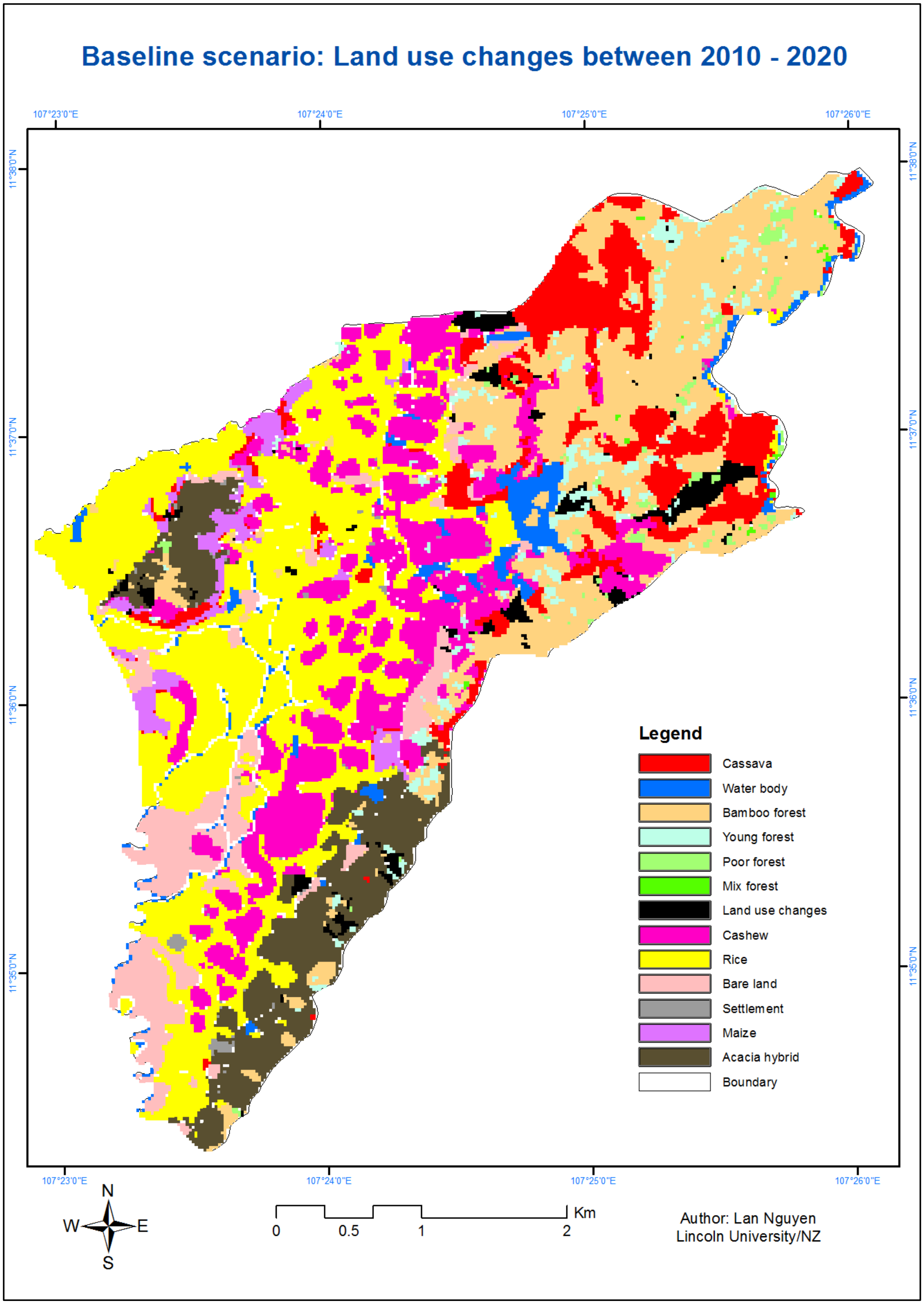
The simulation of the land use changes was initially activated without any policy interventions and parameters were set up to best represent the baseline situation when no development programmes or policies were implemented. The output was the forecasted land use map after a 10-year period starting from 2010. Land use data were overlaid against topographic data to identify potential risks for environmental outcomes such as carbon sequestration loss, soil erosion and forest fragmentation.

5.2.1. Simulation outputs

Parameters of the baseline scenario include local population growth at 2.1%, labour growth at 0.21% (assuming labour force grows 10 times slower than population growth), a current income per day of NZD$7 with 5% annual growth and a payment for forest protection of NZD$20/ha/year for any type of forest cover. There was no payment for forest environment services (PFES) applied yet, and there was no promotion for cashew and acacia hybrid. Those parameters were summarised from different socioeconomic reports issued by local authorities and were also estimated from survey data.

The data were dissolved by household identification (ID) and land use cover (CLU) for both years of 2010 and 2020. The results are illustrated in Figure 25.

Figure 25. Land use land cover of My Lam commune before and after simulation - baseline scenario



As seen in Figure 25, both land use maps of 2010 and 2020 use the same thematic legend to visualise land use allocations and land use changes over the landscape. The land use changed areas were filled with black colour.

Data for each land use cover type and its corresponding area were extracted and put in Table 18 for comparison.

Table 18. Summary of land use land cover before and after simulation

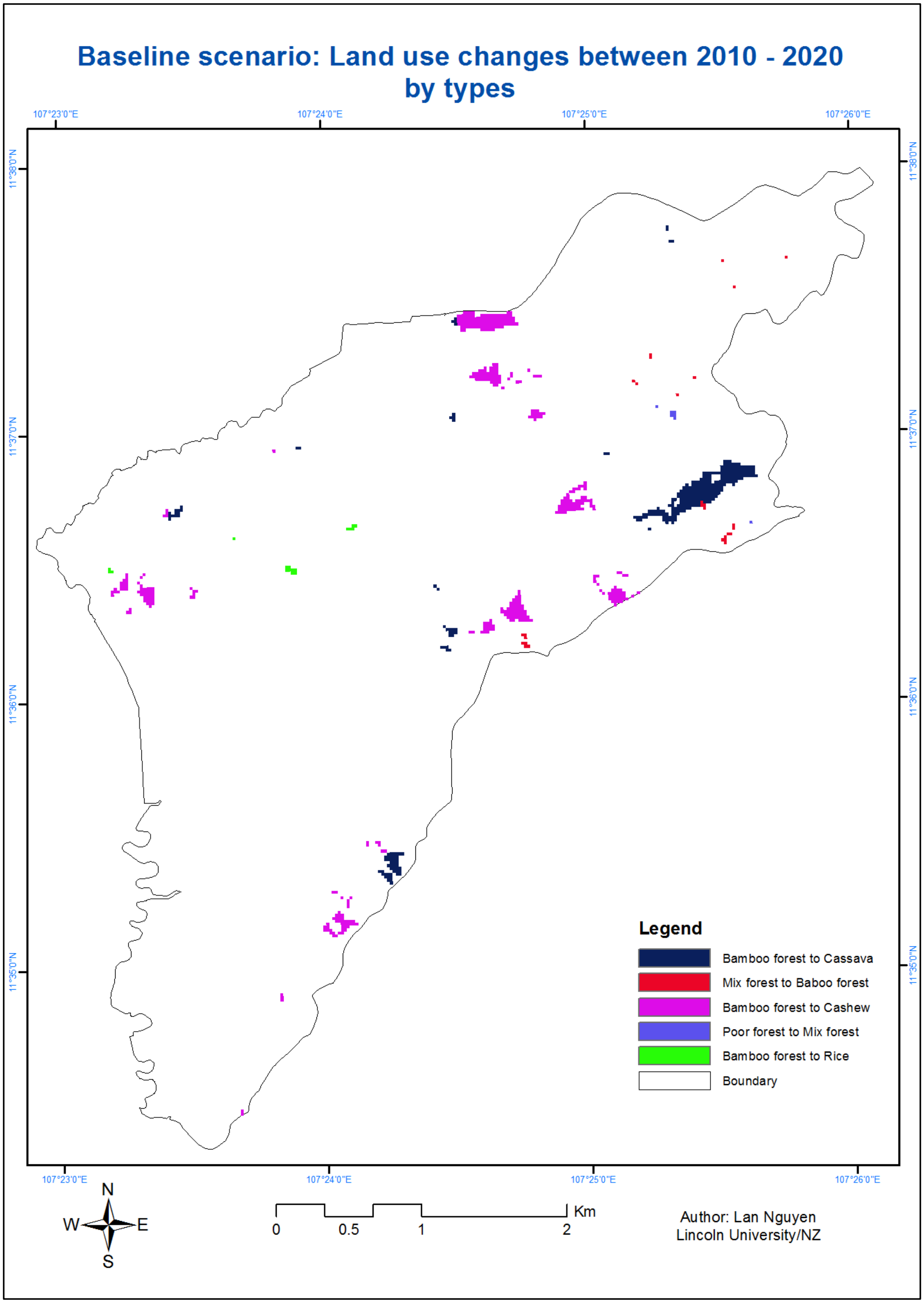
Land use land cover	**Area of 2010 (ha)**	**Area of 2020 (ha)**	**Difference in %**
Poor forest	13	12.72	-2.2
Mixed forest	3.12	2.28	-26.9
Bamboo forest	328.84	296.84	-9.7
Rice field	423.28	423.96	0.2
Bare land	101.08	101.08	0.0
Settlement	7.76	7.76	0.0
Water body	45.84	45.84	0.0
Young forest	42.8	42.8	0.0
Cashew	252.56	271.36	7.4
Acacia	140.44	139.12	-0.9
Maize	37.72	37.72	0.0
Cassava	147.8	162.76	10.1
Total	**1,544.6**	**1,544.6**	

In Table 18 the simulation for the year 2020 showed that, for the baseline scenario, the commune could have more "Cassava", which increased from 147.8 ha in 2010 to 162.76 ha in 2020. Areas of "Poor forest" slightly decreased by 2.2%. "Mixed forest" areas dropped from 3.12ha in 2010 to 2.28ha in 2020 or 26.9%. Areas of "Bamboo forest" decreased by 9.7% in 2020 to 296.84 ha. There was a rise in "Cashew" areas in 2020 by 7.4%. For other land use cover types, no changes were detected. Since the total area of the commune did not change, it can be understood that those changes in areas were just from some land use covers to others.

The information from Table is limited because it only explains the area changes in a quantitative way. There is no visualization of changes over the study site. The image overlaying technique and the confusion matrix are more informative.

Figure 26 visually represents the locations of land use changes in the study area and the details of how land use covers changed from one to other types after 10 years.

Figure 26. Details of land use cover changes from 2010 to 2020



According to Figure 26, the changes happened mostly in the North-East part of commune and by breaking down the changes into categories, it shows that the highest share of change belonged to the transition from "Bamboo forest" to "Cassava". However, other types of forest also had an increasing trend. By converting the land use data of 2010 and 2020 from vector format (polygons) to raster format, it provides a chance to build up the confusion matrix as presented in Table 19.

Table 19. Matrix of land use changes and their areas from 2010 to 2020 in study the area (percentage) - baseline scenario

Land use cover	Poor forest	Mixed forest	Bamboo forest	Rice field	Cashew	Acacia hybrid	Maize	Cassava	Bare land	Settlement	Water body	Young forest	**Total**
Poor forest	97.8	2.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**100**
Mixed forest	0.0	64.1	35.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**100**
Bamboo forest	0.0	0.0	89.9	0.2	5.7	0.0	0.0	4.1	0.0	0.0	0.0	0.0	**100**
Rice field	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**100**
Cashew	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**100**
Acacia hybrid	0.0	0.0	0.0	0.0	0.0	99.1	0.0	0.9	0.0	0.0	0.0	0.0	**100**
Maize	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	**100**
Cassava	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	**100**
Bare land	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	**100**
Settlement	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	**100**
Water body	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	**100**
Young forest	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	**100**

The first row of Table 19 lists the land use cover type of 2020 while its first column shows the land use cover types in 2010. Each cell of the matrix is the conversion in percentages from one land use cover type in 2010 to other types in 2020. For example, only 2.2% of "Poor forest" areas from 2010 were upgraded to "Mixed forest" after 10 years and 97.8% remained as "Poor forest". 39.5% of "Mixed forest" was downgraded to "Bamboo forest" after 10 years. At the same time, 5.7% of "Bamboo forest" were converted to "Cashew" and about 4.1% converted to "Cassava". A small part (0.9%) of "Acacia hybrid" was converted to "Cassava". Land use types such as "Maize", "Bare land", "Resettlement" and "Water body" classes were stable over the simulation period when 100% of them keep their land use cover types in 2020.

5.2.2. Environmental degradation analysis

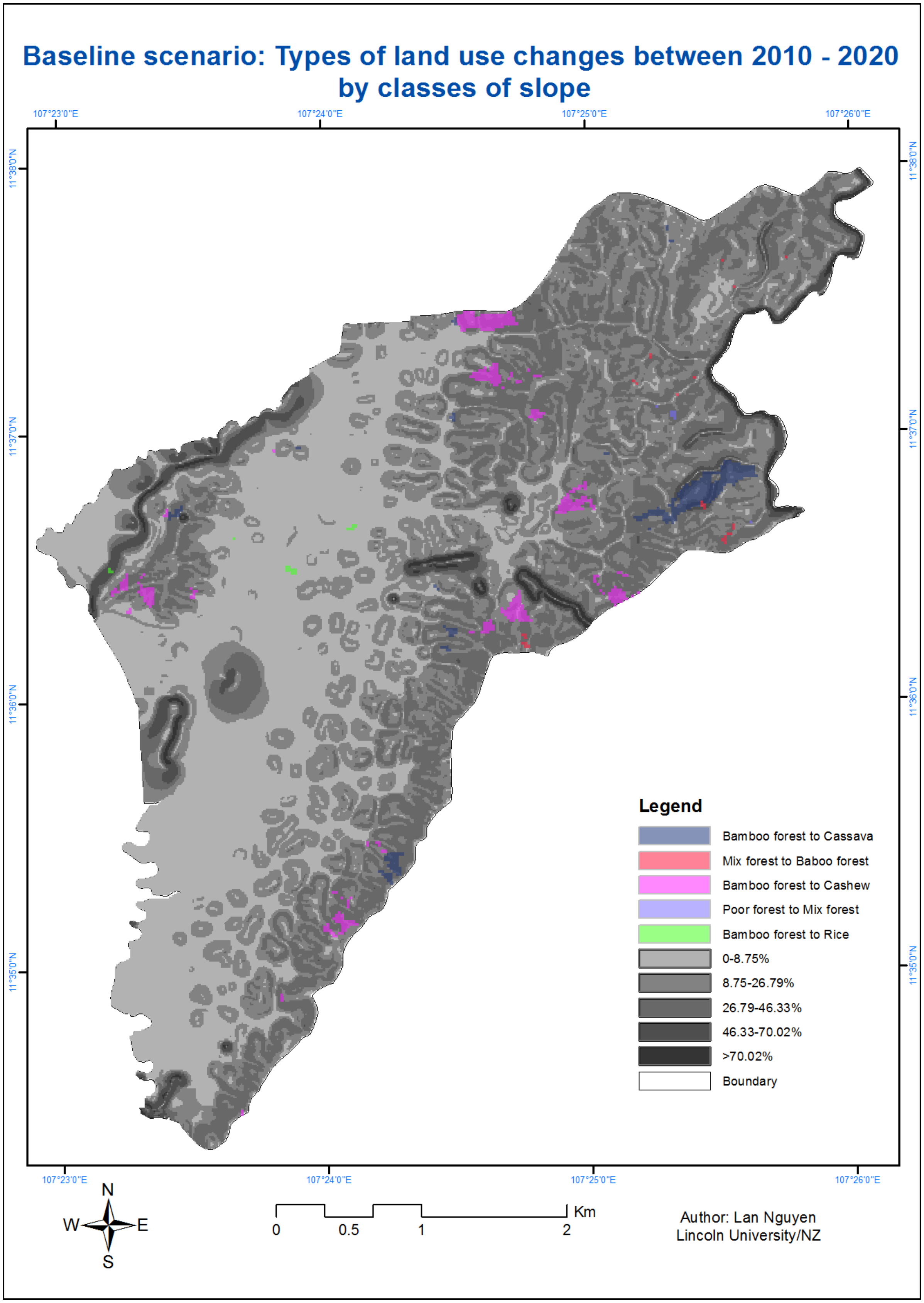
The expected answer for the question about how the future of LULC changes could affect the environmental degradation is the most important part of this chapter. The model uses land use change as a proxy for environmental degradation. There are three kinds of degradation that can be anticipated due to the land use changes: soil erosion, decreased carbon storage and forest fragmentation. The supporting data for those estimations were adapted from other research with an assumption that those data could be applied to the current study area.

5.2.2.1. Soil erosion

LULC change is one negative causal factor affecting the speed of soil erosion. Lack of above ground cover stimulates top soil layers being washed by rain and wind. These losses degrade soil quality and soil structure. The basic input data for erosion estimation was taken from Table 5 in section 3.5.2.

Calculations based on 2010 land use data show that 9,757.5 tonnes of top soil could be washed away per year. The simulated data results for 2020 show a higher amount of washable soil of 10,517.5 tonnes; it has increased approximately by 7.8% after 10 years from the erosion level of 2010.

Figure 27. Distribution of land use changes from 2010 to 2020, by slope classes



By overlaying the areas of land use changes between 2010 and 2020 on the classes of slope raster, the output map in Figure 27 visualises the allocation and distribution of land use changes over the landscape and the slope classes under them. The impression is that land use changes over 10 years happened mainly in areas with steeper slopes.

To quantify the changes of land use covers in each class of slope, two rasters of land use changes and classes of slope were combined and the data of the output raster were tabulated to get Table 20.

Table 20. Changes in soil erosion by class of slope (ton/ha/year) - baseline scenario

	Class 1 0-8.75%	Class 2 8.75-26.79%	Class 3 26.79-46.33%	Class 4 46.33-70.02%	Class 5 >70.02%	Total
From Bamboo forest to Cassava	3.8	62.2	201.2	50.7	0.0	**318.0**
From Mixed forest to Bamboo forest	0.0	0.0	0.0	0.0	0.0	**0.0**
From Bamboo forest to Cashew	8.1	112.6	319.6	0.0	0.0	**440.3**
From Poor forest to Mixed forest	0.0	0.0	0.0	0.0	0.0	**0.0**
From Bamboo forest to Rice	5.3	0.0	0.0	3.6	0.0	**8.9**
Total	**17.2**	**174.9**	**520.8**	**54.4**	**0.0**	**767.2**

Table 20 shows that from 2010 to 2020, land use changes from one type to another caused soil erosion of 767.2 ton/year. Erosion happened mostly in the second and third class of slope at 174.9 ton/year and 520.8 ton/year respectively. That is about 92% of total soil erosion in this scenario. By types of land use change, changing from "Bamboo forest" to "Cashew" and from "Bamboo forest" to "Cassava" caused most of the erosion with 318 ton/year and 440.3 ton/year respectively. Changing from one forest cover type to another forest cover type did not result in much soil erosion. In this scenario, conversion from "Bamboo forest" to "Rice" caused an erosion of 8.9 ton/year across the slope classes.

5.2.2.2. Carbon storage

Carbon storage calculated using vegetative covers could serve as an indicator for environmental quality. Carbon sequestration is estimated based on the data from Table 4 (section 3.4.1).

The total carbon contents are 52,267.7 tonnes and 50,358.7 tonnes for the year 2010 and 2020 respectively. Simulation with baseline parameters suggests a decrease of 1,909 tonnes (or about 3.7%)

after 10 years. These simple carbon content calculations are just snapshots of carbon sequestration in each year based on the areas of land use cover types. There is no way to precisely estimate the amount of carbon accumulated or released throughout the years due to the lack of the qualitative metrics of the forest types. To quantify how the changes in land use covers affect the carbon content of the landscape, net carbon change was calculated and put in Table 21.

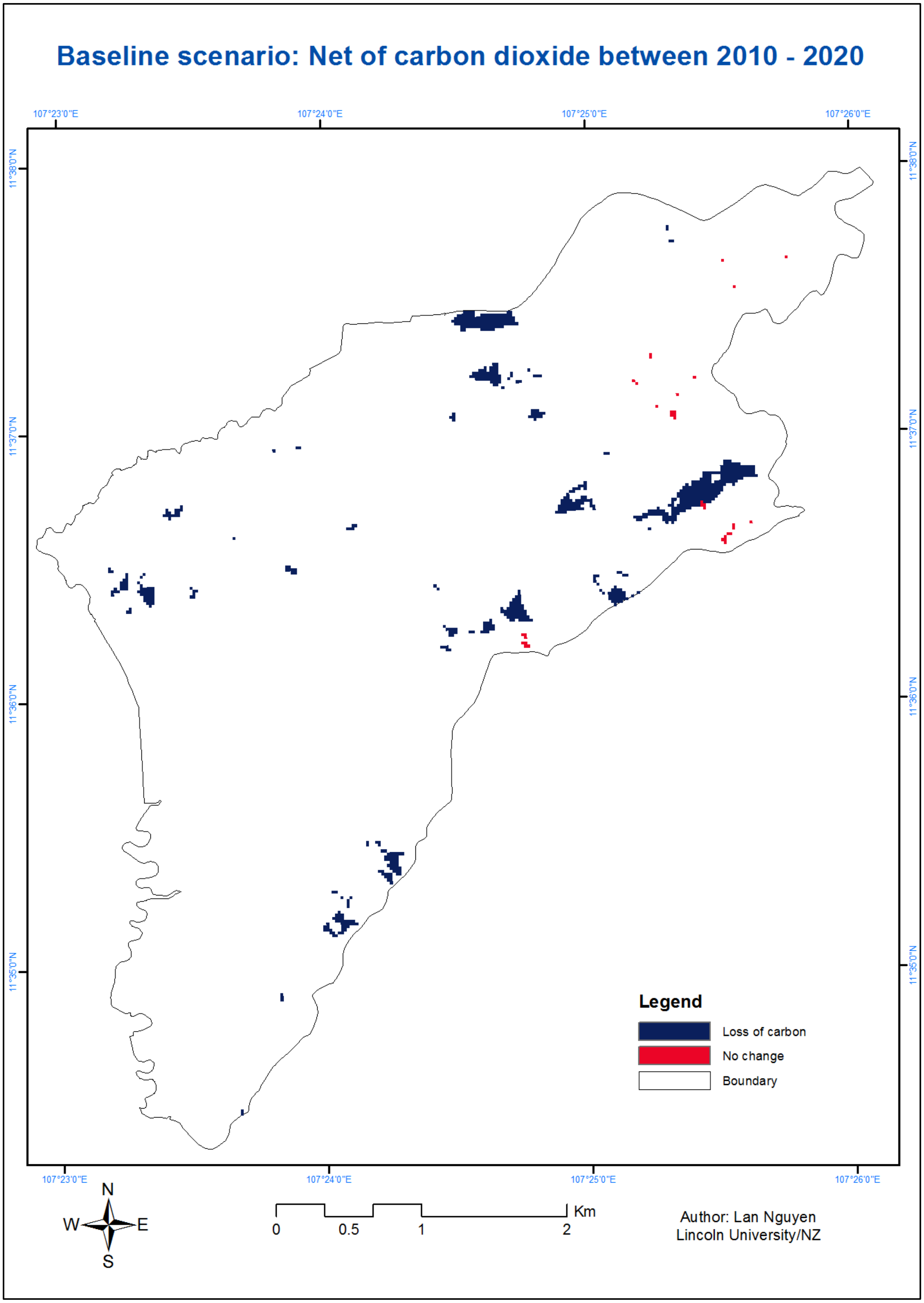
Table 21. Net carbon changes from 2010 to 2020 by land use cover change categories - baseline scenario

	Area (ha)	Carbon change (tonnes)
Bamboo forest to Cassava	13.64	-1118.48
Mix forest to Bamboo forest	1.12	0
Bamboo forest to Cashew	18.64	-689.68
Poor forest to Mix forest	0.28	0
Bamboo forest to Rice	0.68	-55.76
Total	34.36	-1863.92

In Table 21, it is clear that two land use cover changes in the baseline simulation have no gain in carbon content. Changing from "Bamboo forest" to "Cassava", "Cashew" and "Rice" caused a net loss of carbon. The landscape loses 1,118.5 tonnes of carbon after 10 years when "Bamboo forest" was utilised in to "Cassava". Converting to "Cashew" from "Bamboo forest" also contributes a loss of 689.6 tonnes carbon. There was a small portion of "Bamboo forest" converted to "Rice", which is why the loss for this type of change is only 55 tonnes. It is worth noting that there is a small difference between the loss of carbon calculated for the whole study area (1,909) and the loss based on the categories of land use change in the study area (1,863.9). The first one was calculated using the whole land use rasters in 2010 and 2020, while the later one was calculated after breaking down the land use changes into categories. Losing cells due to raster operation in ArcGIS caused the difference in the results in this case.

The combined map in Figure 28 spatially represents the net of carbon content for baseline simulation.

Figure 28. Net carbon content allocation - baseline scenario



As presented in Figure 28, changes having negative net carbon are grouped in the "Carbon loss" group and those with zero net carbon are grouped in "No change". In general, there are almost no carbon gains over the landscape. Interchange among forest covers did not change the carbon content. Carbon losses associated with conversion from "Bamboo forest" are spatially distributed over the northern-east area of the landscape.

5.2.2.3. Patch analysis for the baseline scenario

This part discusses the landscape metrics produced by the Patch Analyst model. The shapefile produced by the simulation was used as input for the Patch Analyst module to calculate metrics for the simulated landscape in 2020, while the input shapefile of the simulation was used to calculate metrics for the original landscape in 2010. In these shapefiles, all polygons (or patches) carrying any forest cover type were kept while the rest were dropped because the aim of this analyst was to examine the fragmentation of forest cover, which has influences on biodiversity, rather than all land use types. Landscape metrics for the baseline scenario are represented in Table 22 below.

Table 22. Patch analysis results - baseline scenario

	Poor forest (3)		Mixed forest (5)		Bamboo forest (6)		Young forest (18)	
	2010	2020	2010	2020	2010	2020	2010	2020
Area-weighted mean shape index	1.3	1.3	1.3	1.3	2.0	2.0	1.4	1.4
Edge density (m/ha)	31.8	33.7	13.0	10.3	342.1	321.0	95.9	104.8
Mean patch size (ha)	0.2	0.2	0.1	0.1	1.2	1.4	0.3	0.3
Number of patches (count)	70.0	67.0	38.0	27.0	272.0	218.0	165.0	165.0
Patch size coefficient of variation (%)	124.0	123.2	66.5	68.1	191.2	182.7	105.1	105.1
Patch size standard deviation (ha)	0.2	0.2	0.1	0.1	2.3	2.5	0.3	0.3
Class area (ha)	12.7	12.4	3.5	2.5	329.7	297.8	42.8	42.8

(3) – land use code *Differences shown in red text*

In Table 22, there are four forest classes joined into this fragmentation analysis: "Poor forest", "Mixed forest", "Bamboo forest" and "Young forest". Simulation with the baseline scenario did not influence the Area-weighted mean shape index and Mean shape index among those forest classes. Those indices which are greater than 1 indicate that patches in all classes have irregular shapes, or are different from a circular shape, which has an index of 1. The more irregular the shapes are, the stronger the likelihood of being naturally formed is (rather than being managed by humans).

The Total edge for "Poor forest", "Mixed forest" and "Bamboo forest" decreased after simulation, indicating that patches in those classes have less complex shapes compared to their shapes in the original landscape. Reducing the number of edges helps patches to reduce the exposure to other types of patches.

The number of patches reduced in "Poor forest", "Mixed forest" and "Bamboo forest" combining with the Class area indicate that there were no new patches generated.

The Patch size standard deviations for "Poor forest", "Mixed forest" and "Young forest" did not change, so the average patch sizes were the same before and after simulation. However, the standard deviation of patch size in "Bamboo forest" slightly changed from 2.3 ha to 2.5 ha. This class may be less fragmented due to the smaller class area and number of patches, while the average size of the patch is bigger.

Although the results show "Patch size coefficients of variation" for the first three classes went down, they are almost the same between 2010 and 2020 in each class, which means the relative variation in patch sizes is at the same level.

Table 22 also shows that the forest classes have a decreasing trend in the baseline scenario where their class areas decreased, and they could be converted into other LULC during simulation. In the attempt to answer the second research question of how some environmental outcomes could potentially be impacted in the future in a "business as usual" context, the SeABM has simulated the land cover changes throughout 10 years starting from 2010 using parameters for a baseline setup to reflect a "no policy intervention" scenario. In each iteration of the simulation, agents or households made their land use decisions considering several constraints and the results of the simulation show that the majority of land use changes happened among the forest cover types. They upgraded from one poorer type to another richer type after several years. This process is similar to what happens in nature. However, there are a few land use changes in the landscape driven by household decision making. The impact of land use changes on environmental degradation is measured using two environmental outcomes: the soil erosion and the carbon content of the landscape. Without any policy intervention, in the projected year the soil erosion increased by 7.8% and carbon content decreased by 3.7% compared to 2010 levels. Although the changes are small in quantity, increased soil erosion carries a negative impact on the landscape. The changes in soil erosion and carbon content were also assessed by their intensity in each type of land use changes and their allocation and distribution on the landscape. "Bamboo forest" seems to be fragmented in 2020, after being converted partially into other types of land use.

5.3. Results for other scenarios

Macro policies were designed to achieve the targets of socioeconomic development at different levels in Vietnam. However, they tend to be applied to administrative entities much larger than the commune level. Policy interventions can be tax reductions or subsidises for specific commodities, subsidies for production and discounted interest rates for groups of producers or beneficiaries with clear intentions to promote the production of certain commodities.

Following this logic, a commune may apply different promotion programmes to target local socioeconomic achievement to match the common theme of a policy of development. Based on the general guidelines, different programmes were interpreted and specified for the commune level to turn a generic instruction into practice. As introduced in section 3.2.6 beside a baseline scenario, six other scenarios were designed to represent six different promotion programmes at the commune level. Those scenarios were fed into simulation to estimate their influence on land use changes.

It is important to note that the six scenarios were investigated independently and any two of the three last scenarios (cashew, acacia hybrid and payment for forest environmental services promotions) cannot be activated simultaneously. However, low population growth, high income and financial support scenarios can be combined with each other and with any of the three last scenarios. Low population growth, high income and financial support scenarios have the same output data and they are similar to the outputs of the baseline scenario. Alternations of the population growth rate, income growth rate and financial support to households have no influence on the outputs of simulations compared to the baseline scenario. The other three scenarios of "Cashew promotion", "Acacia hybrid promotion" and "Payment for forest environmental service" produced different simulation outputs from the baseline scenario simulation. Outputs from simulation with these scenarios are reported as below.

5.3.1. Cashew promotion scenario

In this scenario, all possibilities of LULC changes were converted to cashew plantation.

Table 23 summarizes the changes of different types of land use from 2010 to 2020.

Table 23. Summary of land use land cover before and after simulation - cashew promotion scenario

Land uses	2010	2020	Difference in hectare	In % of 2010
Poor forest	13	13	0	0
Mixed forest	3.12	2	-1.12	-35.9
Bamboo forest	328.84	295.72	-33.12	-10.1
Rice	423.28	423.28	0	0
Bare land	101.08	101.08	0	0
Settlement	7.76	7.76	0	0
Water body	45.84	45.84	0	0
Young forest	42.8	42.8	0	0
Cashew	252.56	273.16	20.6	8.2
Acacia hybrid	140.44	139.12	-1.32	-0.9
Maize	37.72	37.72	0	0
Cassava	147.8	162.76	14.96	10.1
Total	1544.24	1544.24	0	N/A

N/A - not applicable

According to Table 23, areas of "Poor forest" did not change after 10 years in this scenario. There were decreases in areas of "Mixed forest", "Bamboo forest" and "Acacia hybrid", which were 1.12 ha, 33.12 ha and 1.32 ha respectively or 35.9%, 10.1% and 0.9% respectively. "Cashew" and "Cassava" areas increased by 8.2% and 10.1% respectively, thanks to this promotion scenario. These increases resulted from the losses of other land use types during the simulation process.

Figure 29. Land use land cover of My Lam commune before and after simulation - cashew promotion scenario

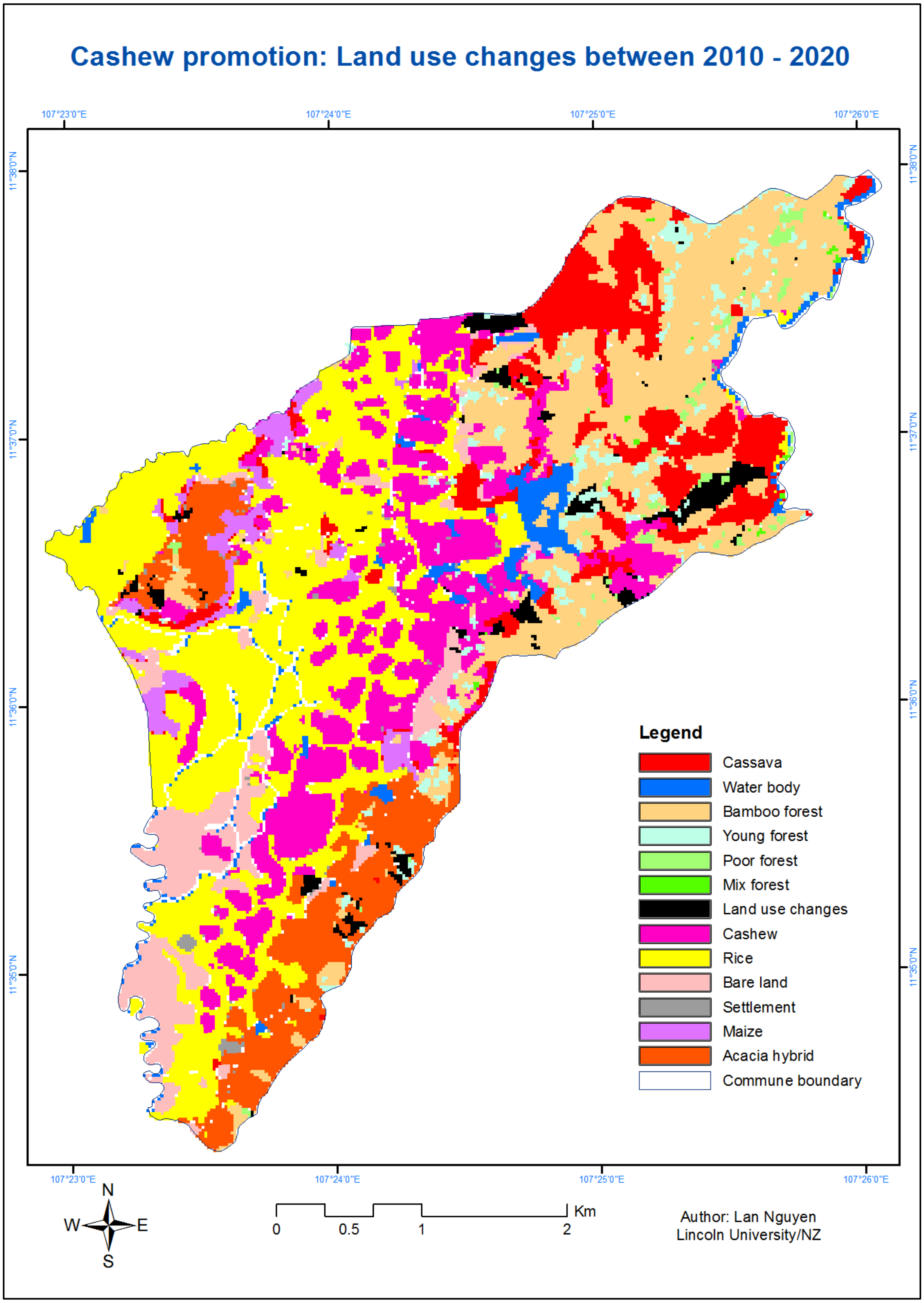


Figure 29 represents the visual distribution of land use change after 10 years under the cashew promotion scenario. However, since the total change is only about 36 ha per 1544 ha (2.3%) it is not clearly shown on the map.

Figure 30. Details of land use cover changes from 2010 to 2020 - cashew promotion scenario

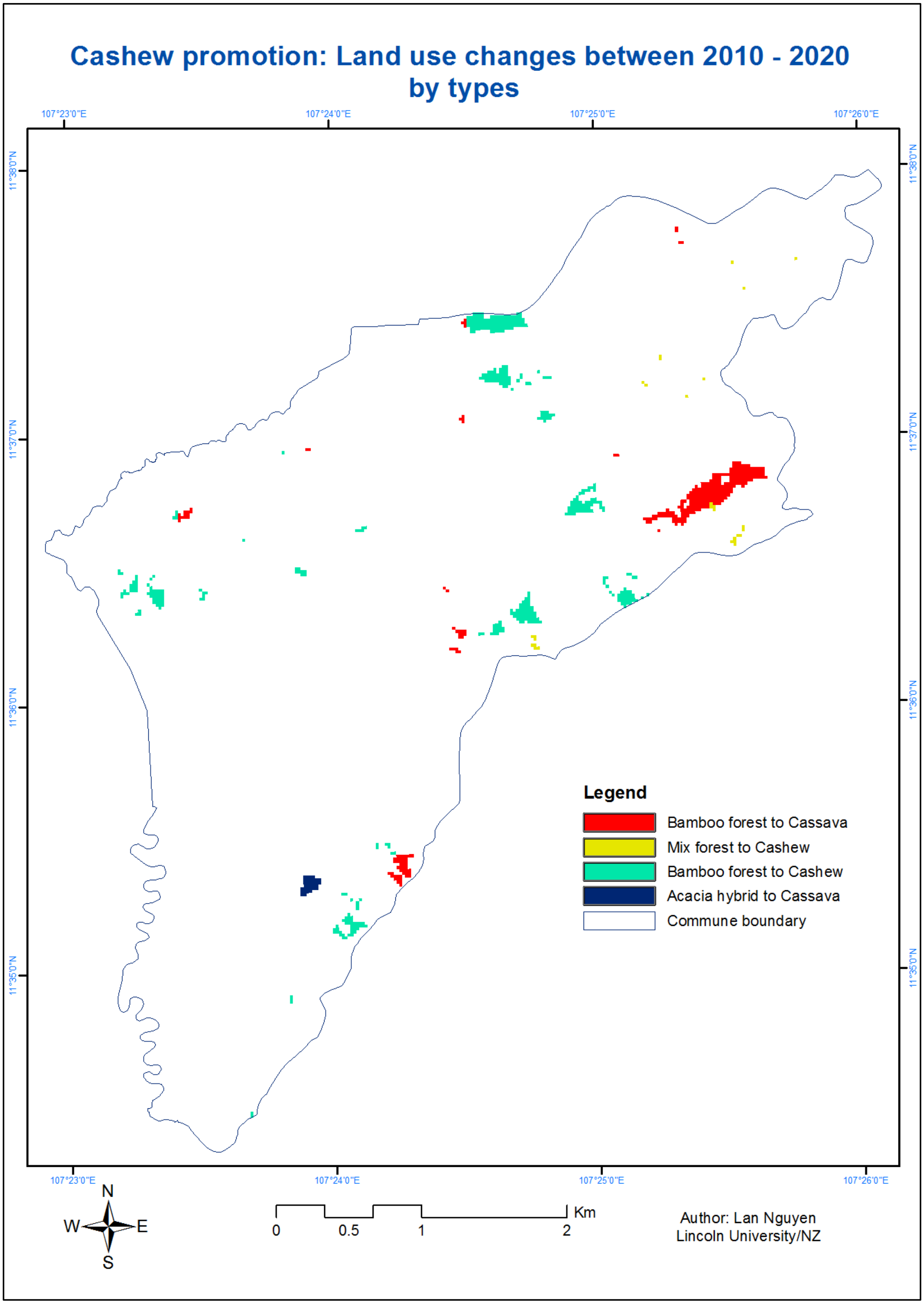


Figure 30 shows the distribution of land use changes over the territory of the study site. There are four major land use changes in this scenario and they distributed in the form of patches over the study area. Popular land use changes seem to be conversions from "Bamboo forest" to "Cashew" and "Cassava".

Table 24. Matrix of land use changes and their areas from 2010 to 2020 in the study area (percentage) - cashew promotion scenario

	Poor forest	Mixed forest	Bamboo forest	Rice	Cashew	Acacia hybrid	Maize	Cassava	Bare land	Settlement	Water body	Young forest	Total
Poor forest	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
Mixed forest	0.0	64.1	0.0	0.0	35.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
Bamboo forest	0.0	0.0	89.9	0.0	5.9	0.0	0.0	4.1	0.0	0.0	0.0	0.0	100.0
Rice	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
Cashew	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
Acacia hybrid	0.0	0.0	0.0	0.0	0.0	99.1	0.0	0.9	0.0	0.0	0.0	0.0	100.0
Maize	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0
Cassava	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	100.0
Bare land	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	100.0
Settlement	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	100.0
Water body	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	100.0
Young forest	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0

Table 24 represents the matrix which was used to quantify the land use changes among different land use cover types from 2010 and 2020. Based on the matrix, 64% of "Mixed forest" area was maintained from 2010 to 2020 and 35.9% converted to "Cashew" by 2020. 89.9% of "Bamboo forest" area was broken down to 5.9% "Cashew" and 4.1% "Cassava". Although the relative figures are small, the absolute figures are about 36 ha in total of original "Bamboo forest" area. A small part of "Acacia hybrid" was converted to "Cassava" after 10 years.

In terms of environmental aspects of land use changes, those changes were overlaid with the classes of slope in the study area to visualize the potential risks from those changes.

Figure 31. Types of land use cover changes of 2010 -2020 by classes of slope - cashew promotion scenario

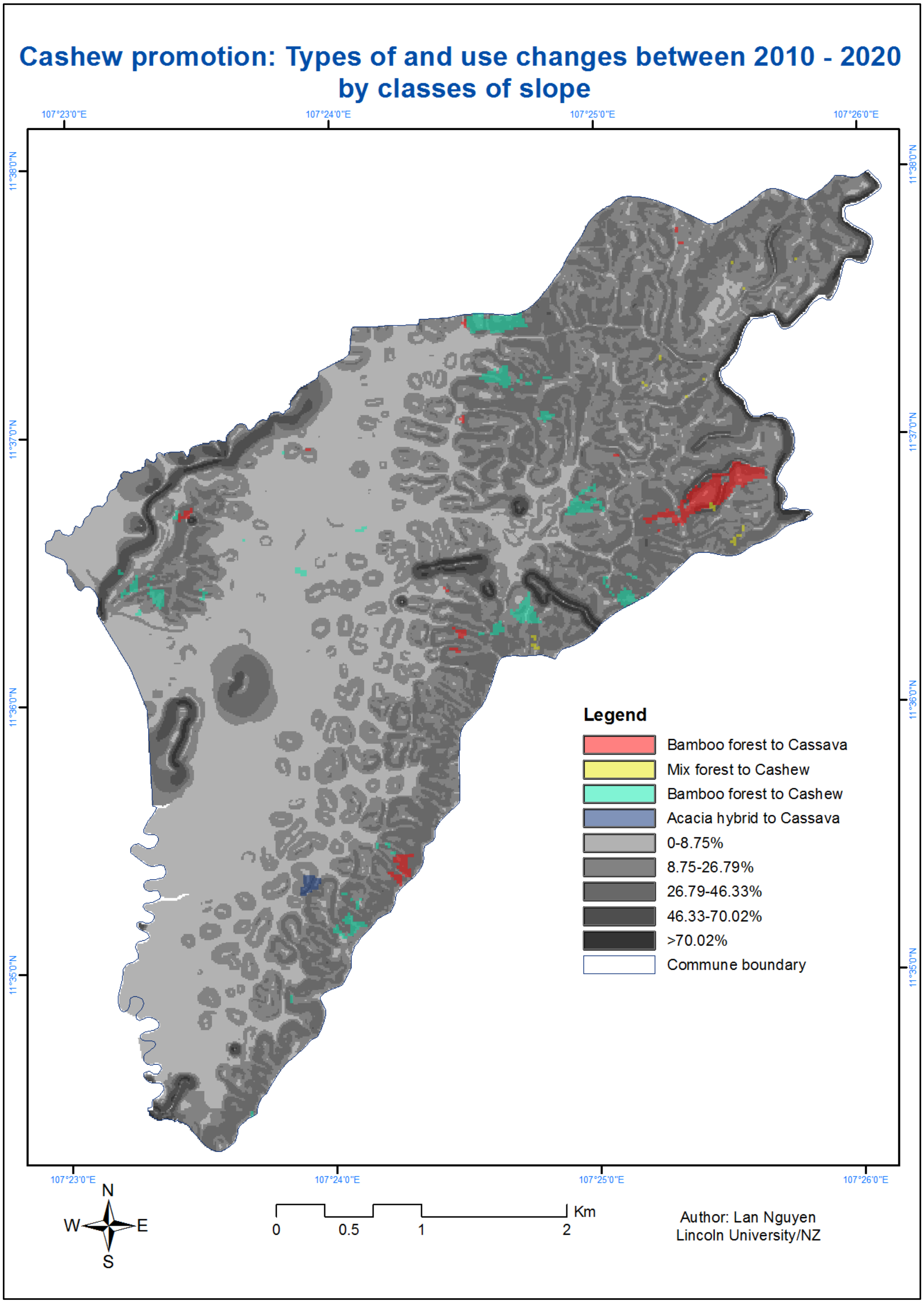


Figure 31 shows that the current scenario caused land use changes in all classes of slope; however, it appears that changes happened more often in areas with slopes between 8.75% and 46.33%.

In Table 25 below, the land use changes are quantified and classified by classes of slope.

Table 25. Area of land use cover changes by class of slope (ha)

Land use cover changes	Class 1 0-8.75%	Class 2 8.75-26.79%	Class 3 26.79-46.33%	Class 4 46.33-70.02%	Class 5 >70.02%	Total
Bamboo forest to Cassava	0.4	4.1	7.5	1.7	0.0	**13.6**
Mixed forest to Cashew	0.0	0.5	0.5	0.1	0.0	**1.1**
Bamboo forest to Cashew	1.1	6.2	11.9	0.1	0.0	**19.3**
Acacia hybrid to Cassava	0.8	0.8	0.2	0.0	0.0	**1.8**
Total	**2.3**	**11.6**	**20.1**	**1.9**	**0.0**	**35.9**

As in Table 25 most land use changes are in Class 2 and Class 3 of slope at 11.6 ha and 20.1 ha respectively. For higher classes of slope there are only small changes of 1.9 ha and 1.8 ha in Class 1 and Class 4 respectively. There were no land use changes in Class 5.

Based on the basic erosion data for vegetation from Table 5 the difference in erosion between types of vegetation and classes of slopes were calculated and fed in to Table 26.

Table 26. Changes in erosion by class of slope (ton/year) - Cashew promotion scenario

	Class 1 0-8.75%	Class 2 8.75-26.79%	Class 3 26.79-46.33%	Class 4 46.33-70.02%	Class 5 >70.02%	Total
Bamboo forest to Cassava	3.8	62.2	201.2	45.2	0.0	**312.4**
Mixed forest to Cashew	0.0	8.7	14.0	3.2	0.0	**25.9**
Bamboo forest to Cashew	16.7	112.6	319.6	3.2	0.0	**452.2**
Acacia hybrid to Cassava	7.2	11.9	6.5	0.0	0.0	**25.6**
Total	**27.7**	**195.4**	**541.2**	**51.6**	**0.0**	**816.1**

In Table 26, changes from "Bamboo forest" to "Cassava" and "Cashew" are 312 ton/year and 452 ton/year respectively. Looking at erosion by class of slope, Class 2 and Class 3 had higher erosion compared to other classes. The highest amount of annual erosion belonged to Class 3 (541 tonnes) and there was no erosion detected in the current scenario for Class 5. Since there is no data related to erosion estimation for "Bare land" and "Settlement", it was not possible to calculate the total annual erosion for the study area. This scenario caused an erosion of 578.2 ton/year in 2020 compared to 2010.

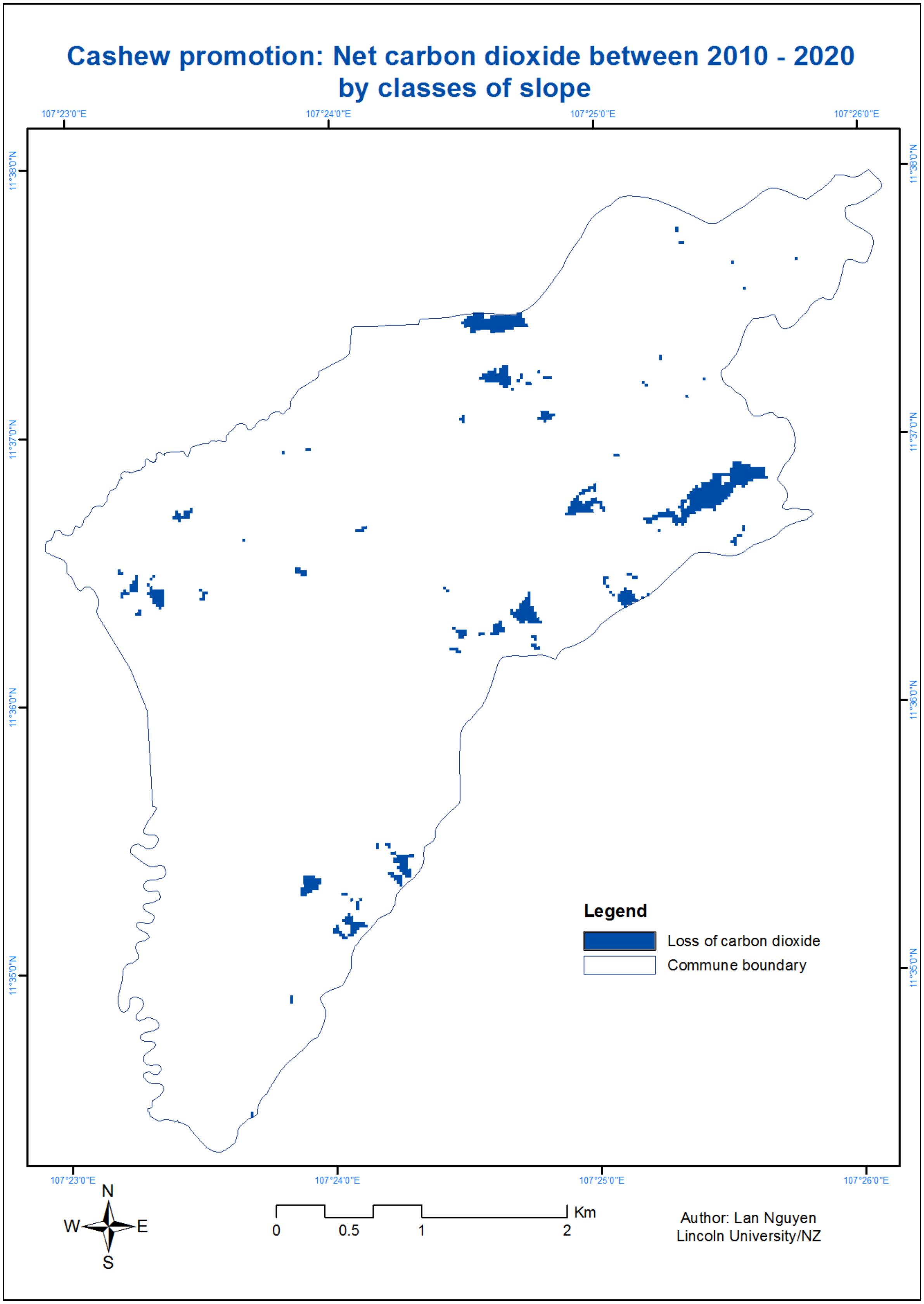
The carbon sequestration was calculated based on data from Table 4. Changes in land use covers, especially in vegetative cover caused changes in carbon storage in the landscape. Based on the basic carbon storage data for each type of forest, the resulting carbon change is represented in Table 27.

Table 27. Net carbon changes from 2010 to 2020 by land use cover change categories - cashew promotion scenario

Types of change	Area	Carbon change (ton/year)
Bamboo forest to Cassava	13.64	-1118.48
Mix forest to Cashew	1.12	-41.44
Bamboo forest to Cashew	19.32	-714.84
Acacia hybrid to Cassava	1.32	-36.96
Total	35.4	-1911.72

As seen in Table 27, for the current scenario, all land use cover changes incurred a loss of carbon. More specifically, 13.64 ha of "Bamboo forest" caused a loss of 1,118.48 tonnes of carbon per year after being converted to "Cassava" and 19.32ha of "Bamboo forest to Cashew" caused a loss of 714.84 tonnes of carbon per year. In total, by pursuing the cashew promotion scenario, the study site could lose up to 1,911.72 tons of carbon per year in 2020. Compared to the carbon storage in 2010 of 52,267.71 tonnes the loss of carbon in 2020 accounts for a loss of 3.7%. The losses of carbon were distributed as represented in Figure 32.

Figure 32. Distribution of net carbon content of 2010-2020 - cashew promotion scenario



In combination with Table 27, Figure 32 represents the spatial distribution of carbon sequestration for the current scenario. There were only carbon losses in this scenario; those losses were distributed over the study area but concentrated in big patches where mainly "Bamboo forest" was converted to other land use covers.

To quantify the fragmentation of landscape, patch analyses at class level were carried out for the vector datasets of 2010 and 2020. In terms of the environmental aspect, "Poor forest", "Bamboo forest", "Mixed forest" and "Young forest" were selected for those analyses. For each class, all calculated indices or landscape metrics are presented in Table 28 for the year 2010 and 2020.

Table 28. Patch analysis results - cashew promotion scenario

	Poor forest (3)		Bamboo forest (6)		Mixed forest (5)		Young forest (18)	
	2010	2020	2010	2020	2010	2020	2010	2020
Area-weighted mean shape index	1.3	1.3	2.0	2.0	1.3	1.3	1.4	1.4
Edge density (m/ha)	31.8	34.8	342.1	324.4	13.0	9.3	95.9	105.2
Mean patch size (ha)	0.2	0.2	1.2	1.4	0.1	0.1	0.3	0.3
Number of patches (count)	70.0	70.0	272.0	217.0	38.0	24.0	165.0	165.0
Patch size coefficient of variation (%)	124.0	124.0	191.2	182.0	66.5	67.3	105.1	105.1
Patch size standard deviation (ha)	0.2	0.2	2.3	2.5	0.1	0.1	0.3	0.3
Class area (ha)	12.7	12.7	329.7	296.6	3.5	2.3	42.8	42.8

(3) - land use code *Differences shown in red text*

Table 28 shows that by applying the Cashew promotion scenario the area-weighted mean shape indices did not change for all classes and they all are greater than 1 or irregular shapes for patches. The more different the shape is from the circular, the less the patch looks managed or created by humans. The changes in values of indices happened mainly with "Bamboo forest" and "Mixed forest" classes. The class areas for "Bamboo forest" and "Mixed forest" reduced in this scenario from 329.7 ha to 296.6 ha and 3.5 ha to 2.3 ha respectively. New areas for cashew plantation had been taken from those two classes but not from other forest coverages like "Poor forest" and "Young forest".

The number of patches also reduced from 272 to 217 for "Bamboo forest" and from 38 to 24 for "Mixed forest". Despite the class area and number of patches being reduced, the mean size of patches in the "Bamboo forest" class increased from 1.2 ha to 1.4 ha at the end of the simulation. The patch size coefficient of variation (%) for "Bamboo forest" also reduced which shows that patch sizes in the simulated landscape did not vary as much as they did in the original 2010 landscape.

The edge density (m/ha) of a class shows the degree of complexity of the edge. Patches with long edges are more exposed to adjacent classes, which may be non-forested cover or may be more isolated and fragmented forest. The increases of edge density in "Poor forest" and "Young forest" show that those classes tend to be fragmented, while decreases for others show that they seem to be less patchy in this scenario.

5.3.2. Acacia hybrid promotion scenario

In this scenario, the simulation framework allowed all suitable land uses to be converted to acacia hybrid. The output of simulation process in 2020 is represented in Table 29.

Table 29. Land use changes from 2010 to 2020 - Acacia hybrid promotion scenario

Land use	2010	2020	Difference in hectare	In % of 2010
Poor forest	13	13	0	0
Mix forest	3.12	2	-1.12	-35.9
Bamboo forest	328.84	295.72	-33.12	-10.1
Rice	423.28	423.28	0	0.0
Bare land	101.08	101.08	0	0.0
Settlement	7.76	7.76	0	0.0
Water body	45.84	45.84	0	0.0
Young forest	42.8	42.8	0	0.0
Cashew	252.56	252.56	0	0.0
Acacia hybrid	140.44	161.04	20.6	14.7
Maize	37.72	37.72	0	0.0
Cassava	147.8	161.44	13.64	9.2
Total	1544.24	1544.24	N/A	N/A

N/A - not applicable

In Table 29 "Bamboo forest" was the main land use cover type to be converted into other land use covers, more than 33 ha or 10% of "Bamboo area" in 2010 was changed to "Acacia hybrid" and "Cassava". This mechanism is similar to the cashew promotion scenario. The area of "Acacia hybrid" was increased by 14.7% according to this scenario.

Figure 33. Distribution of land use cover change from 2010 to 2020 - Acacia hybrid promotion scenario

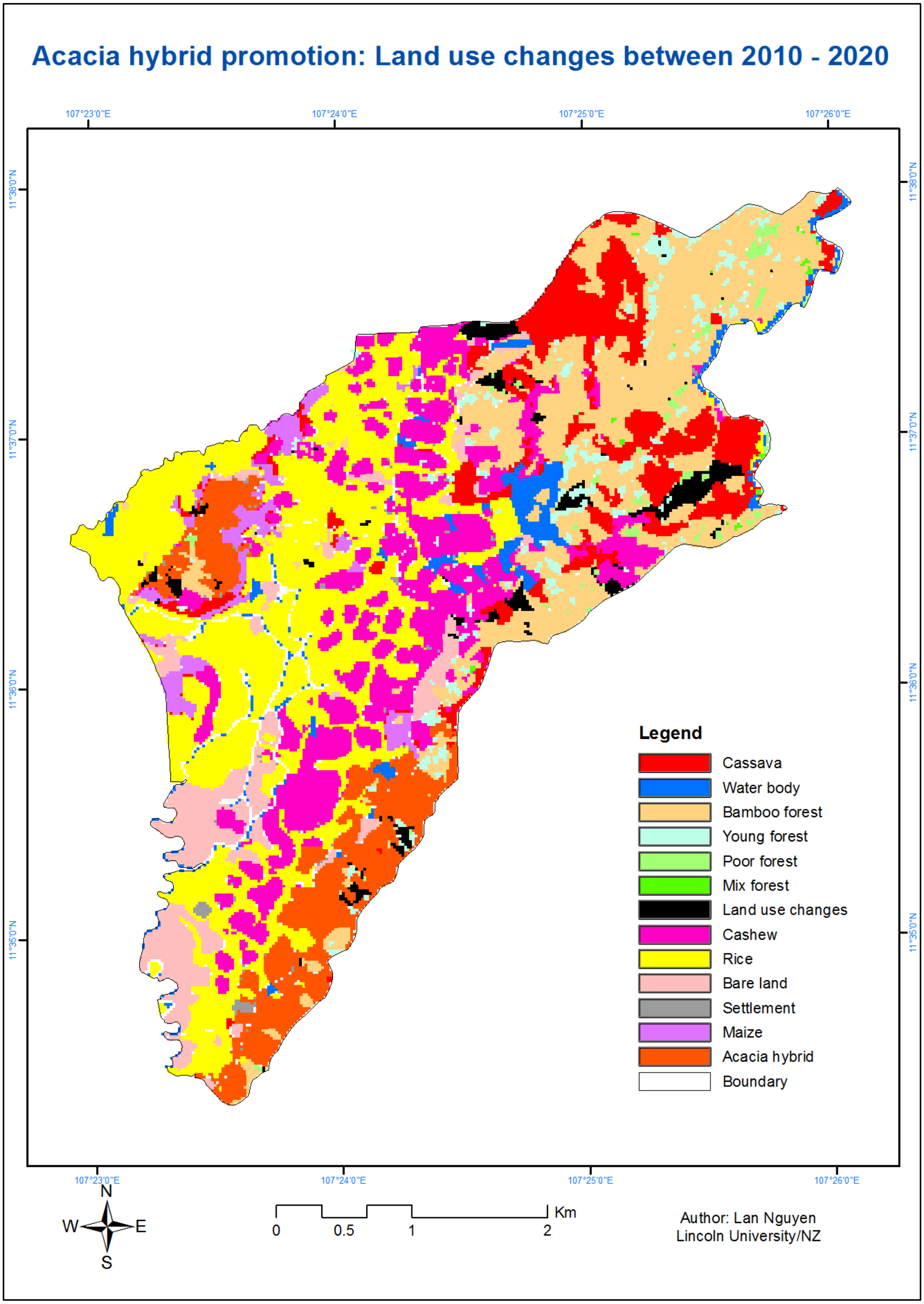
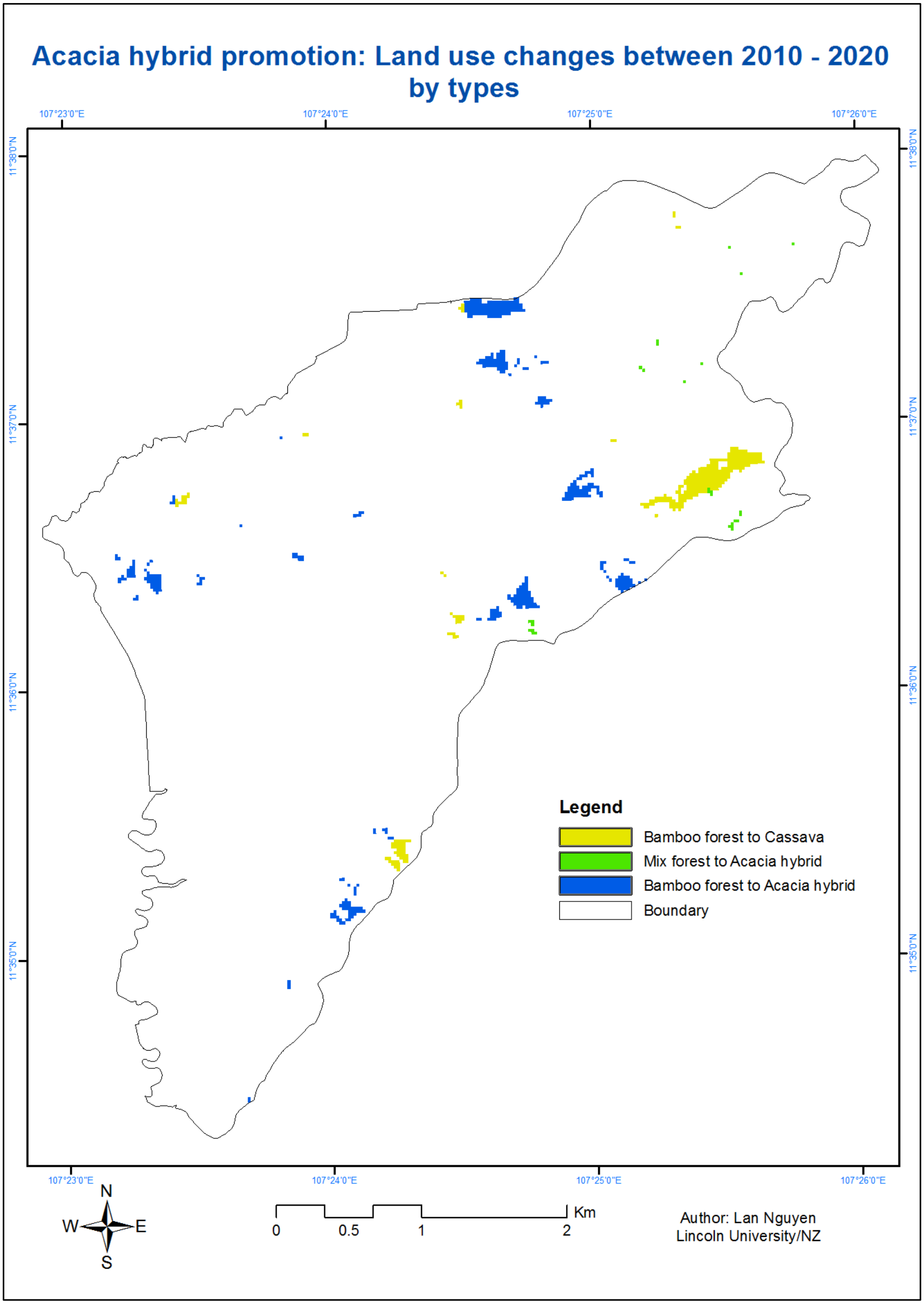


Figure 33 visualises how land use cover changes between 2010 and 2020 were spatially distributed over the study area under the acacia hybrid promotion scenario. It could be identified that the "Bamboo forest" in the eastern side of the study area was changed to other land use covers. However, to quantify the change in both quantity and spatial distribution, map overlaying was applied to extract all changes from one land use cover to others. Figure 34 represents the result of this overlaying technique.

Figure 34. Spatial distribution of land use cover from 2010 to 2020 by types of change



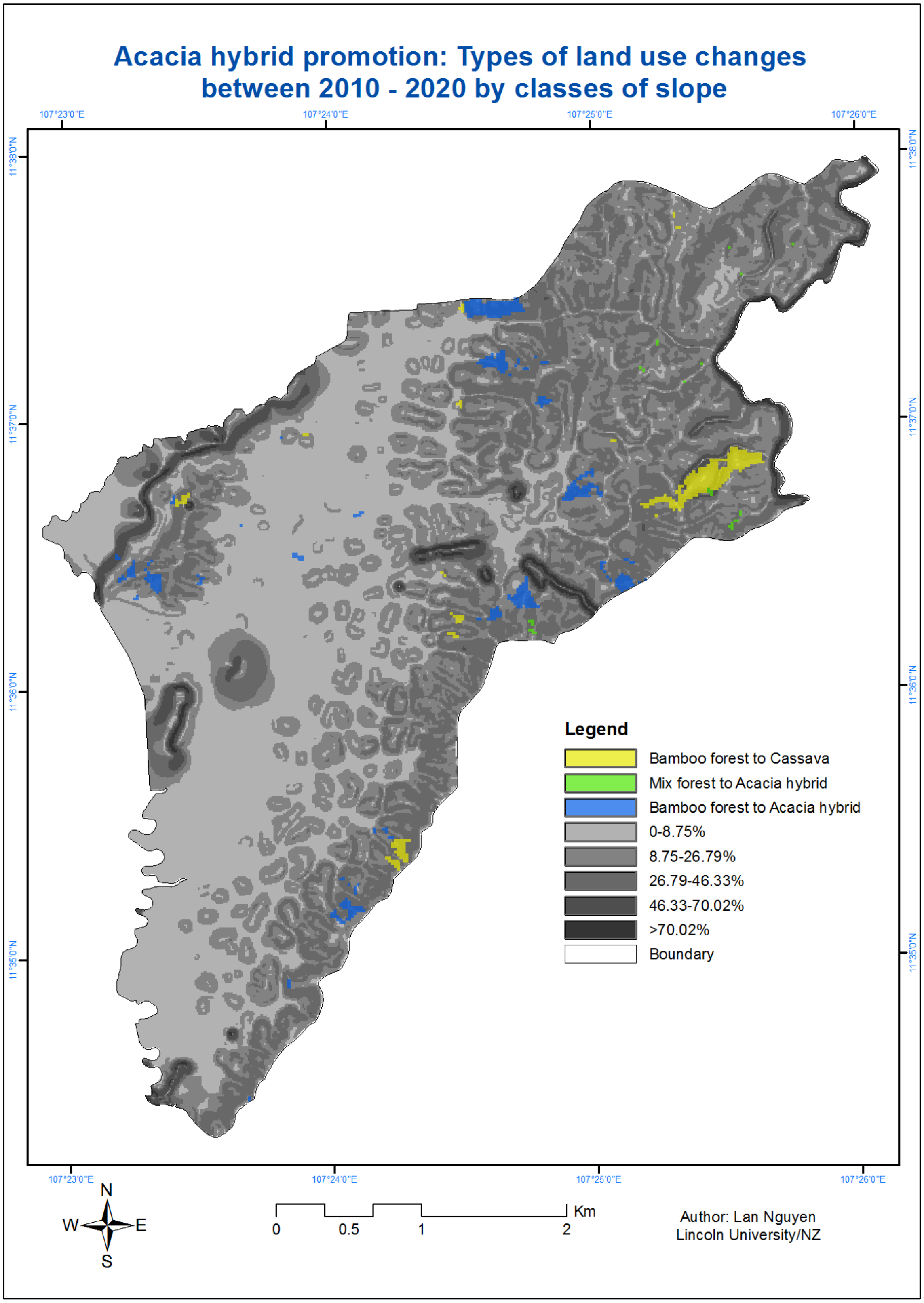
According to the map in Figure 34, there were three popular types of changes in this scenario, which are "Bamboo forest to Cassava" in the eastern side of study area; "Mixed forest to Acacia hybrid" was distributed all over the study area in the form of small patches; and "Bamboo to Acacia hybrid" is mainly in the central part of the study area.

Table 30. Matrix of land use changes and their areas from 2010 to 2020 in the study area (percentage) - Acacia hybrid promotion scenario

	Poor forest	Mixed forest	Bamboo forest	Rice	Cashew	Acacia hybrid	Maize	Cassava	Bare land	Settlement	Water body	Young forest	Total
Poor forest	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
Mixed forest	0.0	64.1	0.0	0.0	0.0	35.9	0.0	0.0	0.0	0.0	0.0	0.0	100.0
Bamboo forest	0.0	0.0	89.9	0.0	0.0	5.9	0.0	4.1	0.0	0.0	0.0	0.0	100.0
Rice	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
Cashew	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
Acacia hybrid	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
Maize	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0
Cassava	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	100.0
Bare land	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	100.0
Settlement	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	100.0
Water body	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	100.0
Young forest	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0

The confusion matrix in Table 30 explains how a land use cover in 2010 could be changed to other land use covers in 2020. This is the result of combining two land use rasters from 2010 and 2020. In this scenario only "Mixed forest" and "Bamboo forest" accounted for the land use cover changes in the 10 yea period. Only 64% of "Mixed forest" area was the same as in 2010 while nearly 36% were converted into "Acacia hybrid". About 90% of "Bamboo forest" areas remained, while about 6% were converted to "Acacia hybrid" and 4% changed to "Cassava" in 2020.

Figure 35. Types of land use cover changes of 2010 -2020 by classes of slope - Acacia hybrid promotion scenario



In Figure 35, the overlaying of types of land use changes with the classes of slope reflects the threat of land slide and soil erosion in the case one land use cover being converted into another land use cover type with a lower resistance to erosion. The spatial distribution shows that changes happened in areas with steep slopes. However, to quantify the amount of potential erosion for the current scenario, data were extracted and processed to present the distribution of types of change by classes of slope as in shown Table 31.

Table 31. Areas of land use cover changes by class of slope (ha) - acacia hybrid promotion scenario

	Class 1 0-8.75%	Class 2 8.75-26.79%	Class 3 26.79-46.33%	Class 4 46.33-70.02%	Class 5 >70.02%	Total
Bamboo forest to Cassava	0.4	4.1	7.5	1.7	0.0	**13.6**
Mixed forest to Acacia hybrid	0.0	0.5	0.5	0.1	0.0	**1.1**
Bamboo forest to Acacia hybrid	1.1	6.2	11.9	0.1	0.0	**19.3**
Total	**1.5**	**10.8**	**19.9**	**1.9**	**0.0**	**34.1**

Table 31 shows that when acacia hybrid is promoted in simulation, land use cover changes happen in Class 2 and Class 3 of slope, where slope varies from 9% to about 47%. It is also clear that there are two types of land use changes dominant in this case, they are "Bamboo forest to Cassava" and "Bamboo forest to Acacia hybrid", with areas of 13.6 ha and 19.3 ha respectively. So, by promoting acacia hybrid, bamboo forest from average slope areas are converted to other land uses, including acacia hybrid. Based on the data from Table 4 in section 3.4.1 combined with data of estimated erosion from different land use covers, Table 32 represents the potential loss of soil by erosion due to the process of land use cover changes in the current scenario for the whole study area.

Table 32. Change in erosion by classes of slope (ton/year) - acacia hybrid promotion scenario

	Class 1 0-8.75%	Class 2 8.75-26.79%	Class 3 26.79-46.33%	Class 4 46.33-70.02%	Class 5 >70.02%	Total
Bamboo forest to Cassava	3.8	62.2	201.2	50.7	0.0	**318.0**
Mixed forest to Acacia hybrid	0.0	0.2	2.0	0.4	0.0	**2.6**
Bamboo forest to Acacia hybrid	0.5	2.2	46.3	0.4	0.0	**49.5**
Total	**4.3**	**64.6**	**249.6**	**51.6**	**0.0**	**370.1**

Data in Table 32 indicates that by promoting acacia hybrid plantation in this scenario, an estimated 370 tonnes of top soil per year could be lost in 2020 compared with the original 2010 landscape. The highest erosion happens when bamboo forest is converted to cassava, mainly in Class 2 of slope. Table 33 summarizes the calculation of carbon content for the current scenario.

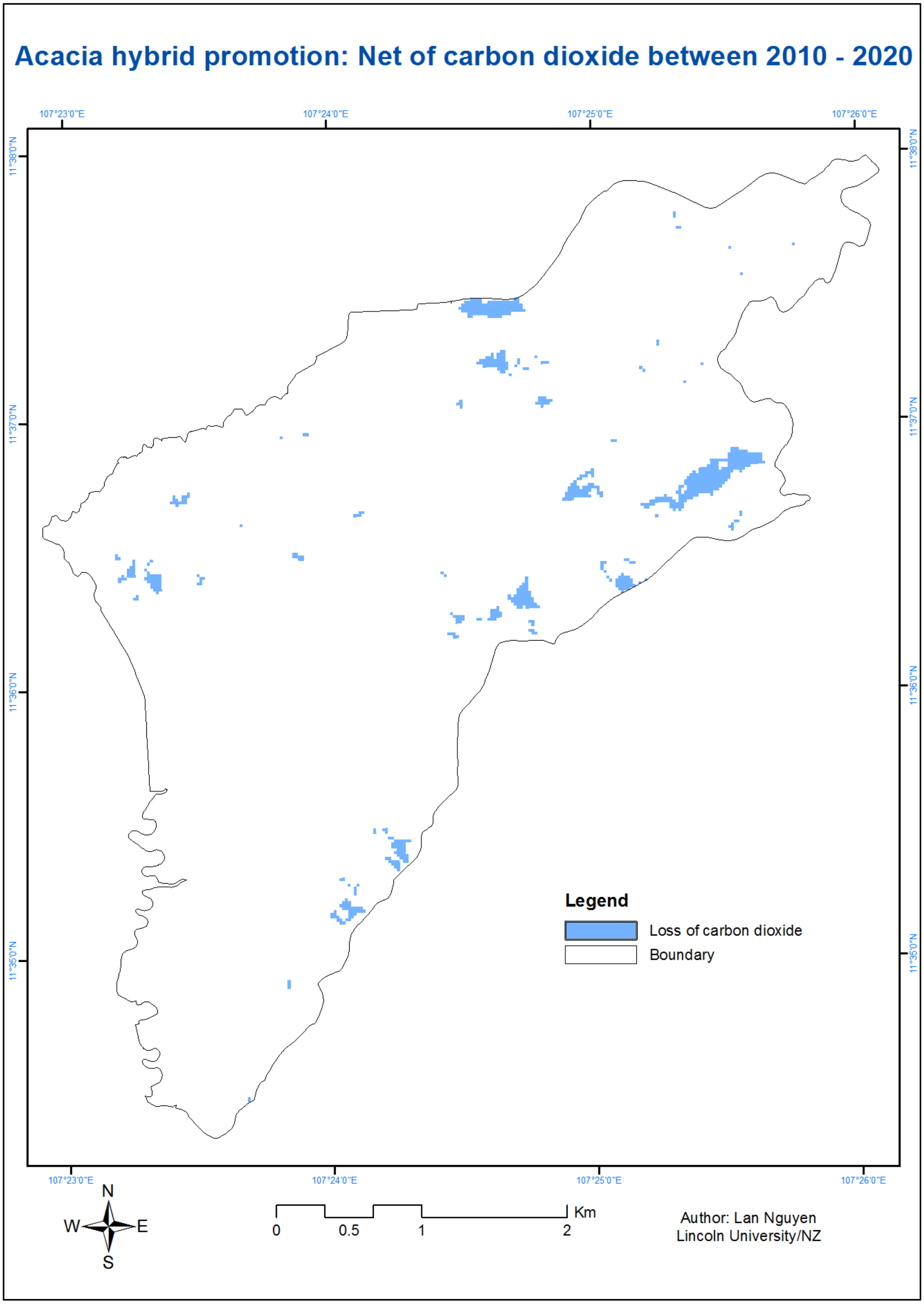
Table 33. Carbon changes by types of land uses (ton/year) - acacia hybrid promotion scenario

Types of change	Area	Carbon change
Bamboo forest to Cassava	13.64	-1118.48
Mixed forest to Acacia hybrid	1.12	-60.48
Bamboo forest to Acacia hybrid	19.32	-1043.28
Total	34.08	-2222.24

As in Table 33, changes from one land use cover to others, especially among forest cover types, greatly affects carbon accumulation in the study site. By applying the acacia hybrid promotion, the study site loses 2,222.24 tonnes of carbon per year in 2020 compared to the original 2010 landscape. The conversion from bamboo forest to cassava and acacia hybrid contributed major parts of the carbon losses (1,118.48 tonnes per year and 1,043.28 tonnes per year) while changing from mixed forest to acacia hybrid only caused a loss of 60.46 tonnes per year.

Figure 36 below shows the spatial distribution of carbon losses over the study area.

Figure 36. Distribution of net carbon content of 2010-2020 - acacia hybrid promotion scenario



Combined with the quantitative data from Table 33, Figure 36 showed that losses mainly happened in bamboo forest areas, in the form of patches and a large part is concentrated in the eastern side of the study area. There is no gain of carbon detected in this scenario.

The patch analysis for the output of simulation was carried out at class level. From the output shapefile of simulated land use at 2020 under the acacia hybrid promotion scenario, four forest cover types were selected to feed the patch analysis. Output indices were generated and are presented in Table 34.

Table 34. Patch analysis results - acacia hybrid promotion scenario

	Poor forest (3)		Mixed forest (5)		Bamboo forest (6)		Young forest (18)	
	2010	2020	2010	2020	2010	2020	2010	2020
Area-weighted mean shape index	1.3	1.3	1.3	1.3	2.0	2.0	1.4	1.4
Edge density (m/ha)	31.8	34.8	13.0	9.3	342.1	324.4	95.9	105.2
Mean patch size (ha)	0.2	0.2	0.1	0.1	1.2	1.4	0.3	0.3
Number of patches (count)	70.0	70.0	38.0	24.0	272.0	217.0	165.0	165.0
Patch size coefficient of variation (%)	124.0	124.0	66.5	67.3	191.2	182.0	105.1	105.1
Patch size standard deviation (ha)	0.2	0.2	0.1	0.1	2.3	2.5	0.3	0.3
Class area (ha)	12.7	12.7	3.5	2.3	329.7	296.6	42.8	42.8

(3) – land use code *Differences shown in red text*

As seen in Table 34, the total area or the sum of all forest cover types, reduced from 388.7 ha to 354.4 ha after 10 years under the promotion of acacia hybrid. However, this reduction results from two forest types: "Bamboo forest" and "Mixed forest". According to above table, "Bamboo forest" occupies 85% of original forested area of 2010 and nearly 84% of the simulated area in 2020.

The mean shape indices for all classes did not change between 2010 and 2020, which indicates that this scenario did not make the shape of patches more or less complicated from standard figures of circles or rectangles.

The number of "Bamboo forest" patches reduced from 272 patches in 2010 to 217 patches in 2020, equivalent to a 20% loss of patches. The average size of "Bamboo forest" patches in 2020 was 1.4 ha while its average size in 2010 was 1.2 ha. The patches in 2020 had lower coefficients of variation (182%) in size compared to their coefficient of variation in 2010 (191.2%). The edge density of "Bamboo forest" class decreased by approximately 5% after 10 years. In this case, "Bamboo forest" patches in the simulated landscape had smaller numbers, bigger sizes and lower edge densities, which could be less patchy than they are in the 2010 original landscape.

The number of "Mixed forest" patches decreased from 38 to 24 after 10 years. Although there was a loss in areas of this class, the average patch size of 0.1 ha for this class did not change after simulation and this indicates that "Mixed forest" patches in the simulated landscape look more isolated than they did in the original landscape.

In terms of landscape fragmentation, the simulated landscape seemed to be less fragmented than the original landscape. Fragmentation in the simulated landscape happened among the "Mixed forest" and "Bamboo forest" patches.

For "Young forest" and "Poor forest" classes, there is not much information to analyse their fragmentation before and after simulation. However, the increases in their edge density make them patchier compared to the original landscape in 2010.

5.3.3. Promotion of payment for forest environmental services (PFES) scenario

In this scenario, a small payment is added to the income of households who have any type of forest in their portfolio. A small change in income is considered in land use decision making by households.

The simulated land use data were tabulated and presented in Table 35.

Table 35. Land use changes from 2010 to 2020 - PFES promotion scenario

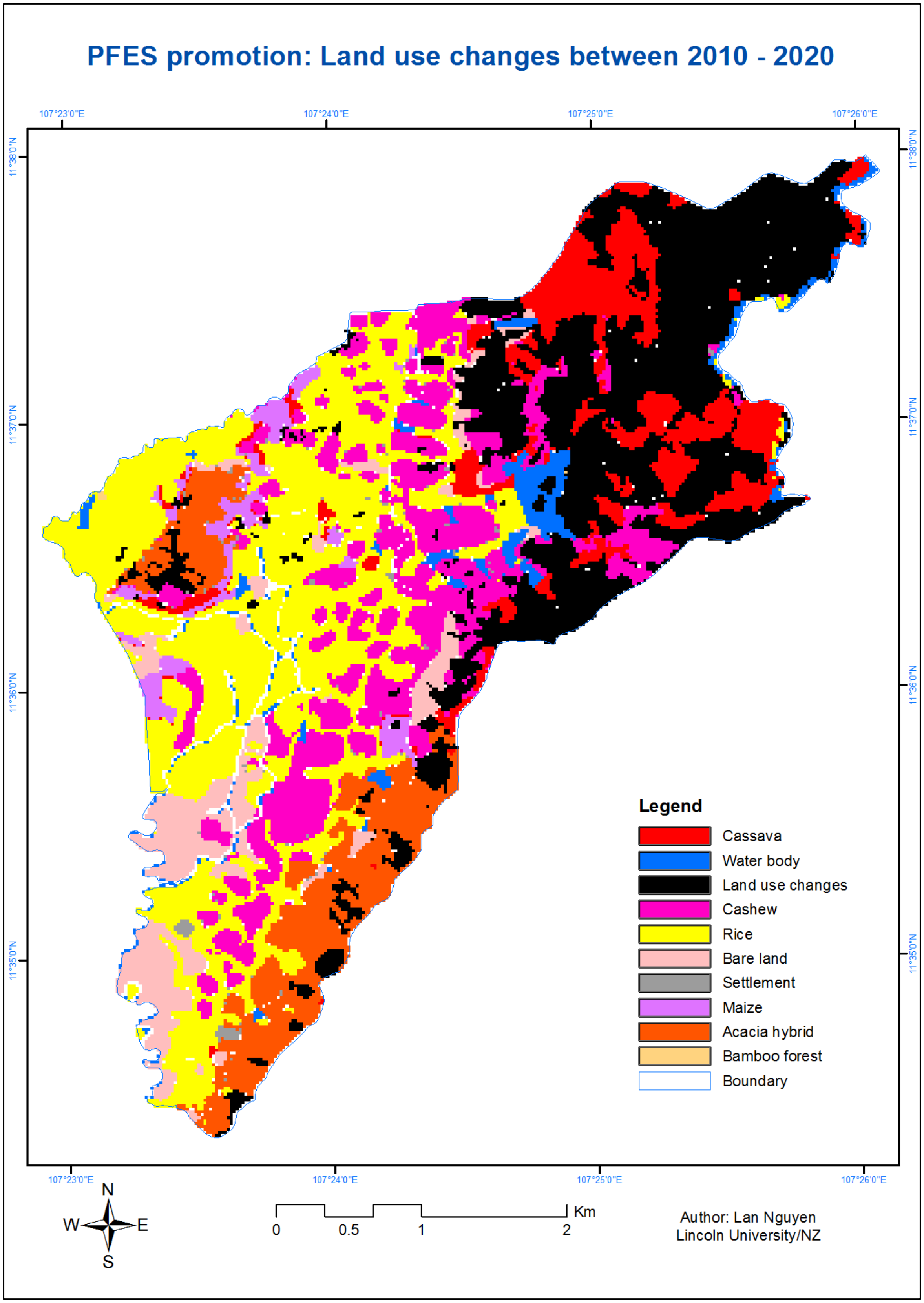
Land use	2010	2020	Difference	In % of 2010
Rich forest	0	3.12	3.12	100
Medium forest	0	13	13.00	100
Poor forest	13	0	-13.00	-100
Mixed forest	3.12	371.28	368.16	11,800
Bamboo forest	328.84	1.12	-327.72	-99.66
Rice	423.28	423.28	0.00	0
Bare land	101.08	101.08	0.00	0
Settlement	7.76	7.76	0.00	0
Water body	45.84	45.84	0.00	0
Young forest	42.8	0	-42.80	-100
Cashew	252.56	251.8	-0.76	-0.30
Acacia hybrid	140.44	140.44	0.00	0
Maize	37.72	37.72	0.00	0
Cassava	147.8	147.8	0.00	0
Total	1544.24	1544.24	N/A	N/A

N/A - not applicable

Data in Table 35 shows that the promotion of payment for forest environmental services hypothesized that households would keep all forest areas and leave them to grow. Under this ideal condition, forest types upgraded to richer types throughout the simulation period. After a 10-year period, a small amount of "Rich forest" and "Medium forest" appeared, and along with that, there was a huge amount of "Mixed forest", which increased from 3.12 ha in 2010 to 371.28 ha. They all resulted from the upgrade of 327.72 ha of "Bamboo forest", 42 ha of "Young forest" and 13 ha of "Poor forest". A very small part of "Cashew" was converted to forested areas. There were no changes detected for other land use cover types in the simulated landscape.

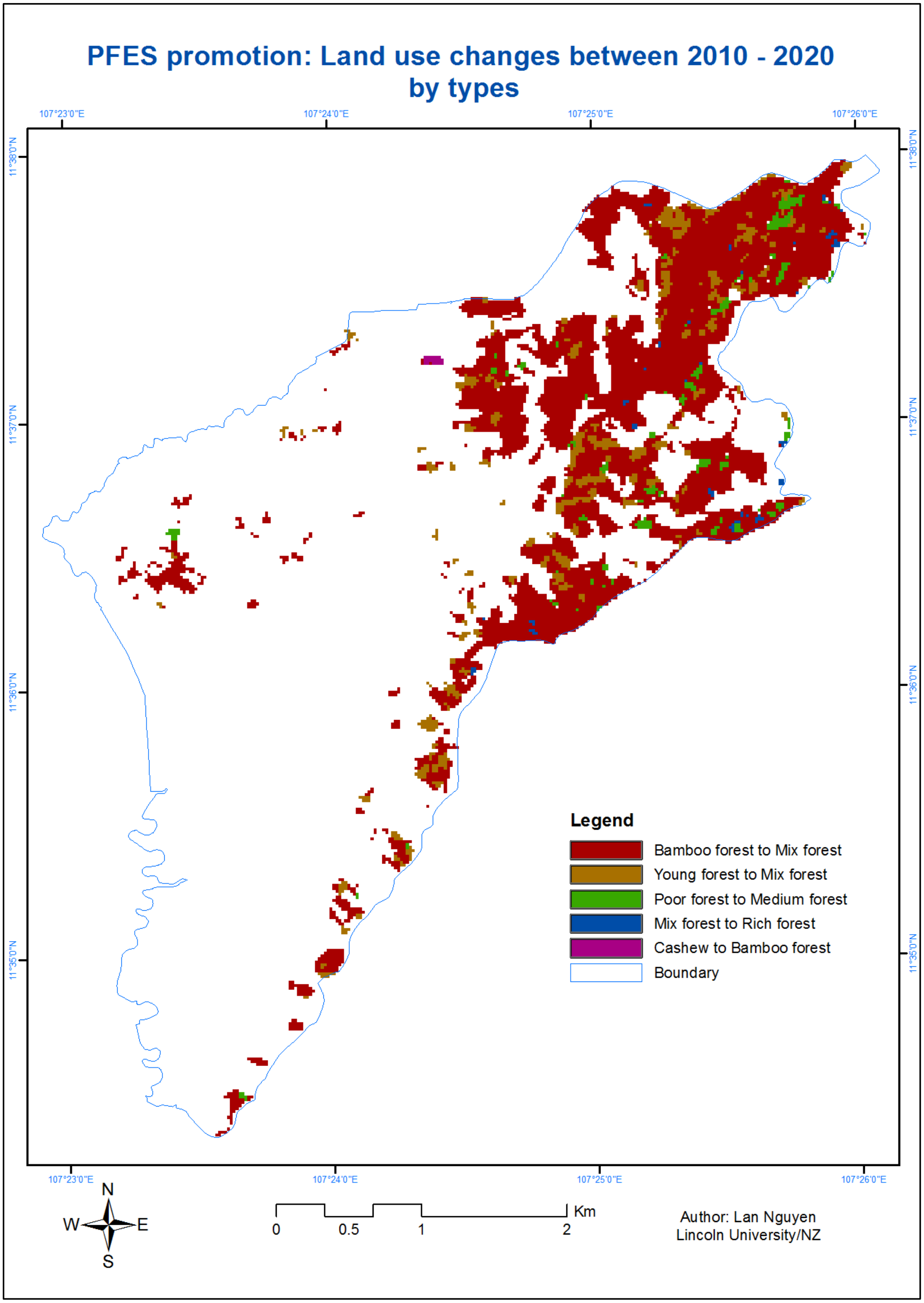
The distribution of quantitative data from Table 35 is presented in Figure 37.

Figure 37. Distribution of land use cover change from 2010 to 2020 - PFES promotion scenario



As seen in Figure 37, since changing from other land use cover to mixed forest is huge, the areas of mixed forest were significant in the land use map of 2020. However, to clearly identify the spatial distribution, the land use cover changes are reflected in Figure 38 below.

Figure 38. Spatial distribution of land use cover change from 2010 to 2020



In Figure 38 the dominant type of land use change in this scenario is "Bamboo forest to Mixed forest", which can be seen in green on the map. This type of change spreads from the north eastern part of the study area down to its southern boundary. There are some small patches of "Young forest" which changed to "Mixed forest". A few spots of "Poor forest" were upgraded to "Medium forest" and other spots of "Mixed forest" grew to "Rich forest". A small change from "Cashew" to "Bamboo forest" was also detected in this spatial distribution.

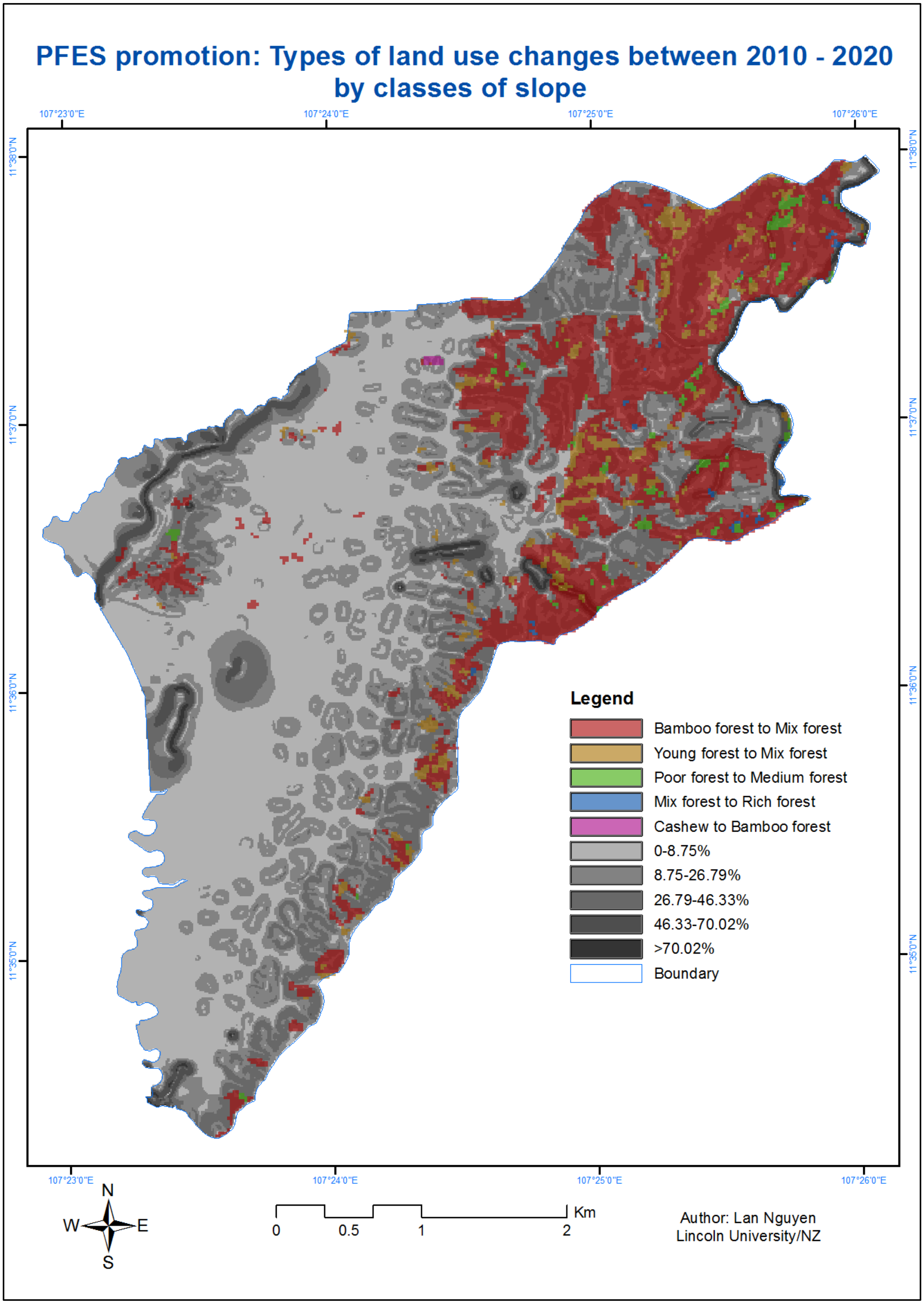
Table 36. Matrix of land use changes and their areas from 2010 to 2020 in the study area (percentage) - PFES promotion scenario

	Rich forest	Medium forest	Poor forest	Mixed forest	Bamboo forest	Rice	Cashew	Acacia hybrid	Maize	Cassava	Bare land	Settle-ment	Water body	Young forest	Total
Rich forest	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
Medium forest	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
Poor forest	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0.0
Mixed forest	100	0	0	0	0	0	0	0	0	0	0	0	0	0	100.0
Bamboo forest	0	0	0	100	0	0	0	0	0	0	0	0	0	0	100.0
Rice	0	0	0	0	0	100	0	0	0	0	0	0	0	0	100.0
Cashew	0	0	0	0	0	0	100	0	0	0	0	0	0	0	100.0
Acacia hybrid	0	0	0	0	0	0	0	100	0	0	0	0	0	0	100.0
Maize	0	0	0	0	0	0	0	0	100	0	0	0	0	0	100.0
Cassava	0	0	0	0	0	0	0	0	0	100	0	0	0	0	100.0
Bare land	0	0	0	0	0	0	0	0	0	0	100	0	0	0	100.0
Settlement	0	0	0	0	0	0	0	0	0	0	0	100	0	0	100.0
Water body	0	0	0	0	0	0	0	0	0	0	0	0	100	0	100.0
Young forest	0	0	0	100	0	0	0	0	0	0	0	0	0	0	100.0

Table 36 represents the confusion matrix aggregated from land use maps of 2010 and 2020 to quantify the change between land use cover types of one year and land use cover types of the other year. According to this matrix, there were no "Rich forest" and "Medium forest" changes from the input dataset of 2010; they all were generated from other land use types due to the configuration of the simulation process under PFES promotion. 100% of "Poor forest" area of 2010 was converted to "Medium forest" in 2020. 100% of "Mixed forest" was upgraded to "Rich forest", and 100% of "Young forest" was upgraded to "Mixed forest" after the simulation.

Figure 39 shows the overlaying of land use cover change by types of change over the classes of slope.

Figure 39. Types of land use cover changes of 2010 -2020 by classes of slope - PFES promotion scenario



The presentation of Figure 39 shows that the changes happened mostly on the hilly areas in the north east and along the southern border of the study area. However, more specific distribution of types of change in each class of slope is calculated in Table 37 below.

Table 37. Area of land use cover changes by class of slope (ha) - PFES promotion scenario

	Class 1 0-8.75%	Class 2 8.75-26.79%	Class 3 26.79-46.33%	Class 4 46.33-70.02%	Class 5 >70.02%	Total
Bamboo forest to Mixed forest	18.9	143.2	154.6	7.9	2.1	**326.8**
Young forest to Mixed forest	3.0	22.1	15.8	1.0	0.8	**42.7**
Poor forest to Medium forest	0.5	3.9	7.1	1.2	0.1	**12.8**
Mixed forest to Rich forest	0.0	1.1	1.6	0.4	0.0	**3.1**
Cashew to Bamboo forest	0.2	0.5	0.0	0.0	0.0	**0.8**
Total	**22.7**	**170.8**	**179.1**	**10.4**	**3.0**	**386.0**

As represented in Table 37, most land use changes under this scenario occurred in the first two classes of slope, 170.8 ha and 179.1 ha were converted in Class 2 and Class 3. At higher slope, only 22.7 ha, 10.4 ha and 3.0 ha were changed in Class 1, 4 and 5 respectively. Total change in the two last classes of slope occupied just less than 10% of the total change of 386 ha. The "Bamboo forest to Mixed forest" was the most dominant type of change with 326.8 ha and, although this type of land use cover change happened at all classes of slope, a significant part was distributed in Class 2 and Class 3. There was a small change of 3.0 ha in Class 5 which was not seen in other scenarios.

Based on the types of changes and areas of each type of changes by classes of slope, the annual erosion in the study area was estimated and reported in Table 38.

Table 38. Change in erosion by classes of slopes (ton/year) - PFES promotion scenario

	Class 1 0-8.75%	Class 2 8.75-26.79%	Class 3 26.79-46.33%	Class 4 46.33-70.02%	Class 5 >70.02%	Total
Bamboo forest to Mixed forest	0.0	0.0	0.0	0.0	0.0	**0.0**
Young forest to Mixed forest	0.0	0.0	0.0	0.0	0.0	**0.0**
Poor forest to Medium forest	0.0	0.0	0.0	0.0	0.0	**0.0**
Mixed forest to Rich forest	0.0	0.0	0.0	0.0	0.0	**0.0**
Cashew to Bamboo forest	-3.72	-9.386	0	0	0	**-13.1**
Total	**-3.7**	**-9.4**	**0.0**	**0.0**	**0.0**	**-13.1**

In Table 38 the estimation shows that by running the simulation with the PFES scenario, the annual soil erosion decreased by 13.1 tonnes in the 2020 landscape. Since 13.1 tonnes of soil erosion per year is very small compared to hundreds of thousands, or even millions of tonnes of soil annually washed from the surface of the study area, it could be considered that erosion the in simulated landscape is not worse than the original 2010 landscape.

The carbon content of the simulated landscape is the next criteria after erosion to qualify the impact of land use cover changes due to the simulation. The result of the carbon content estimation is reported in Table 39.

Table 39. Carbon changes by types of land use change (ton/year) - PFES promotion scenario

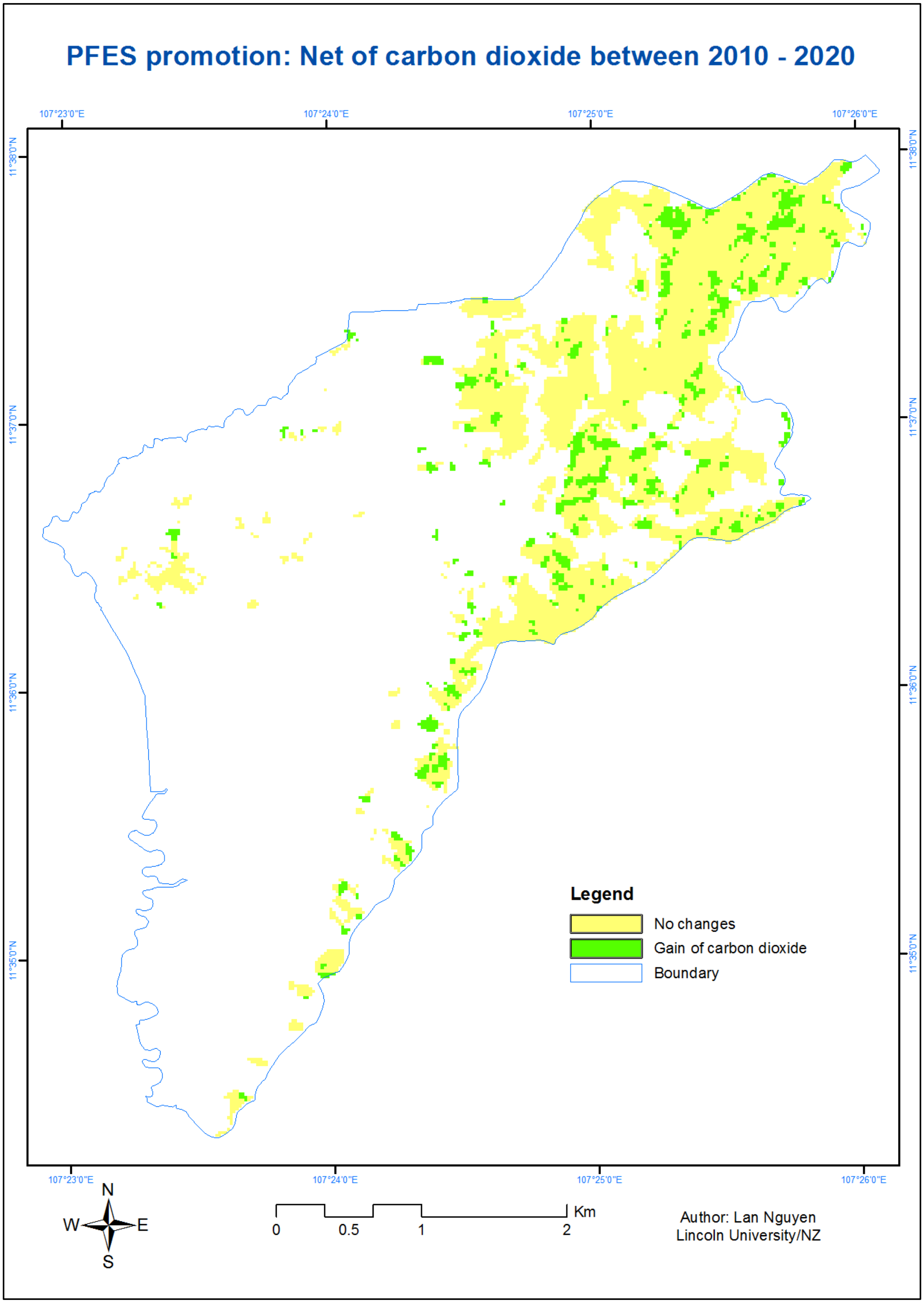
Types of change	Area	Carbon change
Bamboo forest to Mix forest	326.76	0
Young forest to Mix forest	42.68	2219.36
Poor forest to Medium forest	12.76	204.16
Mix forest to Rich forest*	3.08	49.28
Cashew to Bamboo forest	0.76	28.12
Total	386.04	2500.92

**There is no carbon content found for "Rich forest", assume that it is the same as Medium forest, but it could be more than 103 ton/ha/year*

Table 39 shows the PFES scenario promoted forest area to be kept and maintained, and that is why the estimation of carbon content is positive for the simulated landscape, driven by the increase of some forest cover types. Since the reference data for carbon absorption by "Rich forest" is not available, the data of "Medium forest" was taken for calculation. However, "Rich forest" itself may absorb more carbon than "Medium forest" due to the higher standing volume. Based on this assumption, the simulated landscape gained at least 2,500 tonnes of carbon per year comparing to the original 2010 landscape.

Although "Bamboo forest to Mixed forest" was the dominant type of change in this scenario, it does not gain any carbon because bamboo forest and mixed forest seem to have the same carbon absorption level. 42.68ha of young forest changing to mixed forest made a huge difference of 2219.36 tonnes or 88.7% of total carbon gained in 2020. The spatial distribution of quantitative data from Table 39 is shown in Figure 40.

Figure 40. Distribution of net carbon content of 2010-2020 - PFES promotion scenario



As presented in Figure 40, for the current scenario there is no carbon loss. However, the changed areas that did not gain any carbon content is huge. Carbon gained in many small patches of land use changes is spread out from the north, east and central parts of study area, and this gain is also happening along the southern border.

The next criterion of environmental outcomes is landscape fragmentation. In order to estimate it, the patch analysis was run with five (05) forest cover classes from the simulated landscape including "Rich forest", "Medium forest", "Poor forest", "Mixed forest" and "Bamboo forest". On one hand there were no "Rich forest" and "Medium forest" resulting from patch analysis on the 2010 dataset and on the other hand, "Poor forest" and "Young forest" did not appear in the output of patch analysis with the 2020 dataset, so it was not enough information to analyse the fragmentation of those forest covers. Results are represented in Table 40 below.

Table 40. Patch analysis results - PFES promotion scenario

	Rich forest (1)		Medium forest (2)		Poor forest (3)		Mixed forest (5)		Bamboo forest (6)		Young forest (18)	
	2010	2020	2010	2020	2010	2020	2010	2020	2010	2020	2010	2020
Area-weighted mean shape index	N/A	1.3	N/A	1.3	1.3	N/A	1.3	1.7	2.0	1.6	1.4	N/A
Edge density (m/ha)	N/A	0.0	N/A	0.0	31.8	N/A	13.0	0.0	342.1	0.0	95.9	N/A
Mean patch size (ha)	N/A	0.9	N/A	1.8	0.2	N/A	0.1	1.4	1.2	0.4	0.3	N/A
Number of patches (count)	N/A	38.0	N/A	70.0	70.0	N/A	38.0	257.0	272.0	3.0	165.0	N/A
Patch size coefficient of variation (%)	N/A	66.5	N/A	124.0	124.0	N/A	66.5	184.7	191.2	75.3	105.1	N/A
Patch size standard deviation (ha)	N/A	0.60	N/A	2.2	0.2	N/A	0.1	2.7	2.3	0.3	0.3	N/A
Class area (ha)	N/A	3.5	N/A	12.7	12.7	N/A	3.5	372.2	329.7	1.1	42.8	N/A

(3) - land use code *Differences are shown in red texts* *N/A - not applicable*

According to Table 40, for this scenario, the class area of "Medium forest" took over from areas of "Bamboo forest". As a result, the area for "Medium forest" class increased from 3.5 ha to 372.2 ha after 10 years, while the area for "Bamboo forest" class decreased from 329.7 ha to 1.1 ha.

The number of "Medium forest" patches increased from 38 to 257 with a much bigger average patch size of 1.4 ha, instead of 0.1 ha in 2010. Those simulated patches also have a higher coefficient of variation of 184.4% with a very low edge density of 0 m/ha, or least contact with other classes. Combining those indices with the area-weighted mean shape index of 1.7 shows that "Medium forest" class in the 2020 simulated landscape was less patchy than it was in the original 2010 landscape.

All indices of "Bamboo forest" class were reduced, and this class also became less patchy than it was in 2010.

Although there are no available data to elaborate on "Rich forest" and "Medium forest", their appearance in the 2020 landscape enriches the environmental value of the landscape, thanks to their forest quality.

The increase of class area of "Mixed forest" and the decrease in area of "Bamboo forest" class shows the tendency of replacing low quality forest with higher quality forest.

CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS

This research set out to explore the link between environmental degradation and humans' interaction with the landscape in the study area. Households in the research area make land use decisions based on a complex set of factors, which ultimately affect environmental outcomes. Chapter 6 highlights the policy implications of the research outcomes. It shows how policy makers could use the modelling framework to simulate different policy scenarios to evaluate their effect on socio-economic development and their impact on the environment. This chapter also summarises the challenges and gaps in methodology and indicates potential future research directions.

6.1. Implications of research outcomes

Deforestation and forest degradation have become major issues for Vietnam. As presented in Chapter 1, considering the consequences of deforestation and forest degradation, the Vietnamese government has put its efforts into establishing and recovering forest cover since 1992, to recover the environmental functions of forests such as erosion protection and flash flooding prevention, and to improve the livelihood of forest-dependent people.

However, despite the efforts by the state government to slow down the rate of forest degradation, Vietnam still suffers from the loss of its forest cover, especially in many highly forested provinces, including Lam Dong province. Regardless of the restrictions in converting forest land into other purposes, households have been utilising the forest land to meet their needs. Deforestation and forest degradation are the reasons for environmental degradation in Lam Dong province and are driven by forest conversion into other land use practices. This kind of conversion was intensified due to the scarcity of agricultural land coupled with high immigration into Lam Dong province. The Kinh migrants took flat lands which were suitable for paddy rice and perennial crops, while other ethnic groups kept cultivating on steeper slopes. The resulting land use fragmentation breaks the landscape configuration and results in a spatially heterogeneous forest cover, which has consequences for quantity and quality of forest environmental services.

Forest lands are converted into agriculture lands due to the subsistence demand for food and income, that is met from cultivating cash crops such as coffee, rubber, cacao and pepper. These facts drive the need to address the reasons behind the losses in quantity and quality of forests. Households usually act more quickly than the official planning directions from government for many reasons, which is one reason why stopping deforestation and forest degradation seems to be challenging. The role of households in land use decision making is extremely important in the dynamics of forest landscape. Understanding the mechanisms behind those decisions could help policy makers to be more effective in their efforts to introduce sustainable development.

6.1.1. Deforestation and forest degradation in Lam Dong province

The combination of biotic and abiotic factors which affect humans' interaction with the landscape has been hypothesized as the cause for the environmental degradation seen in the rapid socioeconomic development of the Central Highlands region in general, and in Lam Dong province in particular. Farmers make land use decisions based on a complex set of factors, which then affect environmental outcomes. This research aimed to understand how farmers' decision making affects environmental degradation and how different policy interventions might influence degradation. To analyse the problem of how land use decision making relates to environmental outcomes, the research set out to address the following research questions:

- What are the main driving factors for environmental degradation in Lam Dong province?
- What are the potential environmental outcomes with no policy interventions?
- What are the effects of policy interventions on environmental outcomes?

The aim for the first research question was to determine the main driving factors for environmental degradation. This question was important because it defined the factors influencing the land use changes, especially changing from forest land to other land uses. To answer the first research question, a mix of household demographic and spatial factors were collected and analysed.

An econometric model used 15 independently collected or pre-processed variables. They included spatially related variables such as change in land use status, elevation, slope, and distance to residential area, water sources, and transportation; and demographic variables such as households' labour and income. The probit regression showed that the availability of labour, elevation and previous land use change status were significant in determining the likelihood of land use changing from one type to others. The model showed that for any parcel of interest, the land use changes in the past reduced its probability of being changed in the current period, while the availability of labour and the elevation value increased that probability. Forested parcels which had not been changed previously to other purposes may be changed later. The probability of being changed could be higher for parcels located at higher elevation. Parcels of households having high labour availability may be changed more easily. The losses of forest cover from those changes cause environmental degradation.

The results are similar to what had been found by other researchers. As reported by Rounsevell et al. (2010), agent decision behaviour is strongly related to non-spatial information, such as the availability of cash and labour at the household level. Millington et al. (2008) listed a few ranges of potential variables including biophysical metrics (initial land use status, production capacity, location, etc.), economic metrics (profit, market demand, financial sources, etc.) and social metrics (age, agent's world view, neighbourhood, etc.). Castella et al. (2005) explored socioeconomic metrics in agents' behaviour including sources of income (farming or non-farming, subsistence or not subsistence, dependence or not on resource exploitation, etc.), road accessibility (based on Euclidean distance to main road types), wealth indicators, accessibility to credit, schooling, etc. Boundeth et al. (2012) used a dummy binary variable which indicates if the distance from the land parcel of interest is within proximity of 500 meters to the road network to estimate the influence of spatial and socio-economic factors on the probability of land use change.

Although the commodity prices were assumed to influence land use dynamics, it proved difficult to integrate this variable into the framework of the econometric model as there was a lack of market data to inform any relevant analysis. However, the effect of commodity prices has been integrated into the profit from agricultural activities in each household. This profit variable was not statistically significant in regression against the dependent variable.

By answering the first research question, this research has sought to understand the influence of driving factors on land use change and, more importantly, the answer has provided a core function for SeABM to create a simulation framework to forecast the land use dynamic from 2010 to 2020 by modelling the interaction of driving factors with the decision-making process.

The purpose of the second research question was to forecast environmental degradation with no changes in policies. The output of the baseline scenario simulation served as input for other models to evaluate the potential erosion, carbon sequestration and fragmentation of this future landscape in 2020. These three environmental outcomes signal the severity of environmental degradation caused by land use change decisions resulting from the baseline scenario.

The baseline scenario was considered as "business as usual" development which reduced forest areas in the future landscape which, in turn, increased the erosion, decreased the carbon sequestration and promoted the fragmentation of forests. The baseline scenario caused a decrease in forest areas in the study area, with "Mixed forest" area reduced by 27% and "Bamboo forest" losing about 10% of its area after 10 years. Those areas were converted to cashew and cassava. This reflects the gradual trend of changing to other high value crops in this area. The decrease in forested area affected the quality of the environment. The amount of washable soil increased by 7.8% and carbon sequestration in the forest decreased by 3.4% in 2020. The patch analysis showed that the losses in area caused fragmentation in "Poor forest", "Mixed forest" and "Bamboo forest" areas. Outputs of the baseline scenario imply that maintaining the current policy with high population growth rate and low income would only lead to more environmental degradation due to the losses of forest for other purposes.

The answer to the second research question showed that in the long term, if the same policy framework (as described in parameters for the baseline scenario) is to be applied, the quality of the environment in the buffer zone of Cat Tien National Park would be degraded. And sooner or later, it would negatively influence the conservation efforts in the core zone of the national park. The current policy framework has a weakness in controlling population growth rate of 2.7% per year. At this rate, the number of mouths to feed in each household and the labour will increase, the cash balance of households may go into deficit and lead to changes in their land uses to compensate for the loss. The land use changes followed the past trend which was described by several land use change modules generated from analysing time-series satellite images. The baseline scenario did not show the effectiveness in reducing forest conversion or improving environmental protection.

The third research question has been addressed through simulations with different policy scenarios representing different development alternatives to compare against the baseline scenario. The alternations of population growth rate, income growth rate and financial support for households produced similar outputs to the baseline scenario. These three scenarios led to forest losses and environmental degradation. However, other policy scenarios of promoting high value crops such as cashew, acacia hybrid and subsidising households with a payment for forest environmental services scheme had positive influences on the outputs of simulations for 2020.

The promotion of cash crops assumed the enthusiastic support of authorities in introducing high value crops to boost the local economy. Promoting cashew accelerated the conversion from forest land to cashew plantation and cassava. About 36% of "Mixed forest" areas were lost in this scenario and, as expected from the losses of forest land, the environment was degraded. The amount of washable soil increased by 8.4% in 2020 with the erosion happening mostly on moderate slopes (26% to 46%). The losses of forest also decreased the carbon sequestration in the simulated landscape by 3.7%. Cashew is a perennial crop but due to its characteristics and the low density of planting, it causes erosion if it is planted on slopes without a low vegetative cover to protect topsoil and it cannot sequester carbon dioxide as well as ordinary forests do. And in this case, converting forest land to other non-forestry purposes increased the fragmentation of forest patches.

The acacia hybrid promotion scenario was similar to the cashew promotion when suitable areas for perennial crops were converted to acacia hybrid and existing acacia hybrid areas were kept. In the output, the area of acacia hybrid and cassava increased by approximately 15% and 10% respectively. The amount of washable soil in this scenario was 3.8% higher than the 2010 landscape. This figure is less than half of what cashew promotion scenario caused. Promoting acacia hybrids decreased the carbon storage by 4.3% compared to 2010 landscape. The short-term production purpose of acacia hybrids excluded it from the forest plantation for environmental improvement, which is why the patch analysis did not include this type of land cover. Losing forests for other purposes, the landscape in 2020 looked more fragmented than in 2010. A high density of acacia hybrid plantations would prevent soil erosion; however its short rotation reduces the ability to absorb carbon dioxide.

The introduction of payment for forest environmental services (PFES) scenario assumed that households who manage forests get a subsidy for not clearing forest. This scenario shifted about 100% of "Bamboo forest" to "Mixed forest". This is the only scenario where "Rich forest" appeared. The improvement in both forest quantity and quality had contributed to the environmental outcomes. Although the amount of washable soil did not reduce much, it did not increase to make the situation worse as it did other scenarios. By keeping and promoting forest area, the carbon sequestration increased by 5% compared to the 2010 landscape. Obviously, the payment for keeping and promoting forest had a positive effect on environmental outcomes. If households could maintain their livelihood from different sustainable sources of income, they would not utilise the forest land for other purposes. In this way, deforestation and forest degradation would be reduced. However, collecting the payment from end users and redistributing it fairly among forest owners would pose its own challenges.

The answers to the second and the third research questions have suggested that the effects of different policy interventions on environmental outcomes has been successfully modelled by comparison of different input parameters on the outputs of simulations. In these cases, the quality of environmental outcomes has been driven by alternating the development policies for the study area.

6.1.2. Research findings

The unique feature of this research is the synergy between various type of data and the dynamic spatial simulation model. The extensive technical work on the design and implementation of the simulation ABM, as well as collecting data including a qualitative survey, helped inform how households make decisions on land uses in such a data-scarce region as the selected case study. This research has enriched the ABM approach in studying land use, especially in exploring and reasoning how farmers' decision making would drive environmental degradation. Simulation using different scenarios helped identify policy options that reduce negative impacts on environment and can be extended to apply to other parts of Vietnam.

This research used an econometric model to incorporate spatial and non-spatial factors into a decision-making function and deployed a hybrid coupling GIS with ABM, which offered a flexible way to represent environmental outcomes distributed spatially over the landscape. This exposed the superiority of geospatial analysis compared to other non-spatial land use modelling approaches.

It is interesting to note that although the Central Highlands region of Vietnam is famous for coffee plantations, people in this commune were not successful with coffee plantation in the past. The reason could be that the micro climate of this commune is hotter and dryer than many other communes/districts in Lam Dong province. Due to this fact, there was no "coffee promotion" scenario designed for the simulation.

The analysis of land use trajectories from temporal satellite images exposed a frequent land use dynamic in this study area from 1995 to 2010. “Medium forest” and “Mixed forest” were detected on the 1995 image when the intensity of human activities in this area was not that high. However, under the pressure of socio-economic development, the high-quality forests were replaced by other poorer land covers, such as “Bamboo forest”, “Poor forest” or were converted to agricultural land. The analysis also showed that after being converted for agricultural purposes, some areas were interchanged among different annual crops due to attempts to find the best crops, and non-preferable areas were abandoned to turn back to “Bamboo forest”. After 15 years, “Medium forest” had largely disappeared from the landscape, which was a sign of degraded environment.

Studying the land use trajectories was a useful approach to understand the trends of land uses under the influence of different factors. Different changes were also associated with other spatial factors such as distance to water or the road network, or slope at the position of change. Based on the historical land use changes, predictions were made for potential land use changes in the future.

The intention of this research was to investigate the influence of land use changes, especially the loss of forest land, on environmental quality, so it had to link how the land use dynamic could degrade the environment at the study area. To simplify the process, the vegetative coverage (except rice and other annual crops) was used as inputs for other models to calculate the potential erosion, carbon sequestration and fragmentation degree of the landscape. Those are environmental outcomes which could characterize the environmental quality and are dependent on vegetative coverage. The models were based on other studies and simplified by using interpolation or extrapolation for missing values to adapt to the current research. The erosion model is more complicated than the carbon sequestration model because it was driven by types of vegetation and slope classes. Each vegetative type has a different ability to hold the runoff on different slopes. The carbon sequestration model is simply based on the area of different vegetative types.

The Patch Analyst extension for ArcGIS was developed based on FRAGSTATS (McGarigal and Marks 1995) with inputs in this case being patches of different forest types. The result offered a range of outputs to measure the fragmentation of the landscape. However, only seven of them were used in this research. Patch Analyst was not only used to estimate the degree of fragmentation in one landscape but also to compare among different scenarios.

According to the simulation results for the 10-year period, “Medium forest” would only appear back on the landscape when applying the payment for forest environmental services scenario, or, in other words, in this simulation framework, forest quality would be improved if households had a sustainable source of income for maintaining forest and its environmental services.

6.2 Key drivers and policy options

An important objective of the research was to provide policy makers useful information in managing land use change and impacts on the environment. The results show that there are useful measures that can be used by policy makers. This includes the assessment of land use change pace; management of land use allocation regarding the elevation and the crops; and the distribution of labour force for agricultural activities. The findings of this research help in identifying some policy implications which may influence future land use decision making practices. The simulations with the scenarios have demonstrated how different interpretations of policies may affect the future landscape. From this point, the policy options can be made to ensure the changes in the landscape follow a desirable direction which satisfies both socio-economic growth and environmental quality. Based on the results of this research, a range of policies would be suggested:

a) Maximising land use conversion if it is not profitable:

The outputs of the econometric model showed that land use is less likely to be changed by households in a period if it has been changed in the previous period. It is good to encourage households to keep certain land use types which are profitable for them and try some new practices on less profitable land uses. However, policy makers should be aware of what types of land use can be converted and converted to what. A case study needs to be set up to prove the benefit of new crops and new cultivation methods to persuade households to intensify agriculture.

Households may make better decisions on land use if the information flow is enhanced. At the commune level, the interpretation of top-down policies is slow and vague. It takes time to translate and adapt the idea of any policy from the upper level down to operative programs because it needs to go through a hierarchy of documentation. And when the interpretation comes out, it is often too late to reflect current circumstances. Households have access to different kinds of market information through TV and radio. However, they do not have access to the good market distribution nodes. Their products have to go through a busy network of middle men in the area and prices become skewed. Households do not really take the initiative in negotiation, so they must diversify cash crops to cope with uncertain market information and try to keep rice land for food security. This is a gap where policy can intervene. Instead of being passive in interacting with the market, operative programs or plans at the commune level should be more robust and connect local production to other potential markets or buyers rather than depending on middle men. Optimizing the supply chain and improving the accessibility for transportation will reduce transaction costs. Households can only intensify agriculture if they have a better orientation.

b) Encouraging long-term investment and long rotation for plantations at high elevation and reducing the restriction on rice land on flat areas:

In this research, the positive relation of elevation on land use change showed that land use change may happen more often at higher elevations. During the survey and data collection at the study area, it was found that most residents in this commune are immigrants from the northern provinces of Vietnam who came and settled in the early 1990s. Flat and low lands were cleared for paddy rice, as they used to cultivate in the north for food security in early settlement. Because all good and flat land had been set aside for rice, other crops had no chance to compete, so they had to be planted in steeper areas. This fact explains the positive sign for the coefficient of elevation factor in the regression model. It indicates that the high spots have a strong correlation with the land use change dynamics in this area. According to the spatial data, the maximum elevation of this area is about 400 m, so except for the low and flat areas for paddy rice, areas up to 400 m were employed for other crops.

The more often the land use happens at higher elevation, the more the land surface is exposed to soil erosion because the root system and canopy coverage take time to form and may not reach the maturity necessary to reduce erosion before being changed to other crops. So, crops or vegetative types for high elevations should be long-term, long rotation to make sure they protect the soil from being washed. Forest cultivation or long rotation acacia could be the potential candidates to solve this issue. The observation and the discussion with households revealed that local farmers just tried to utilise the areas at high elevation to get whatever they can, rather than invest into those areas. Households were dependent on the rice areas at low and flat land and could not change them to other purposes due to the complicated regulations and culture of growing paddy rice. Land use policy needs to be flexible for households to use their flat land to grow anything more profitable than rice while waiting for the long rotation crops or forest plantation at high elevation, which will contribute significantly to reducing environmental degradation.

c) Controlling the birth rate and improving education:

As pointed out by the econometric model, the higher number of labour dedicated for agricultural activities in each household causes the higher probability of land use being changed. Land use changes are not always negative for the environment, however. The surplus of labour in each household allows them to convert more land to housing and cultivation. The access to paddy rice areas is limited so it is a threat to forest lands to be converted. For long-term development, it is necessary for local authorities to control the population growth because land resource is getting scarcer.

The scenario analysis showed that the low natural population growth scenario of 1.5% did not affect the outcomes of land use changes over a period of 10 years for the current study area. However, low population growth should be promoted for sustainable socio-economic development and this is well within the common policy of government. Along with the low natural growth rate, there should be a solution to improve the local economy by maintaining the right size of the labour force for agricultural activities.

Results from the demographic analysis showed that children in the study area start helping their families with a few agricultural activities even before they reach age 15. Young people should be encouraged to have vocational training or higher education to transform the local economy, rather than being dependent on primitive agricultural cultivation. They also need to be introduced to successful business models which can be applied at a local scale and increase value of local products. If the local economy is not attractive enough for young people to stay, then they must still have the skills to find stable and adequate jobs in other cities, rather than being seasonal workers there.

d) Increasing incentive for forest protection:

The Payment for Forest Ecosystem Services (PFES) had been tried in Lam Dong province as a solution to improve forest protection. Payments are based on formal documentation of land allocation/contracts, so to get paid, suppliers of environmental services must hold legal land tenures of their land. Current Land law secures many rights for land users or households; however, the government has ultimate power over the land, including the right to recall and plan land resources according to the requirement of socio-economic development and for national defence purposes. The current leases of 20 years for agricultural land and 50 years for forest land do not encourage land users to put in any long-term investment. This urges a policy amendment to secure long-term ownership for land users.

Reducing Emissions from Deforestation and Forest Degradation (REDD+) has the potential to simultaneously contribute to climate change mitigation and poverty alleviation by generating income when reducing emissions. The Provincial REDD+ Action Plan (PRAP) of Lam Dong province played an important role in creating an adaptive path for emissions reduction through sustainable forest management plans for forestry companies and integrating emissions reduction in master plans for forest protection programmes. New guidance and regulations are based on Ministry of Agriculture and Rural Development of Vietnam (MARD) decrees and will orient the forestry sector towards more sustainable forestry and land use plans. Since a large area of forest is leased to households for cultivation and protection, emission reduction policies should give a fair reward for their efforts to keep and improve the quality of forest. Moreover, those policies recommend that the private sector (including households) invest in low emissions development in the Agriculture, Forestry and Other Land Use (AFOLU) sectors in Lam Dong province.

e) Improving legal responsibility for violation in forest protection:

Local authorities seem reluctant to arrest or prosecute small scale deforestation. It is important to put a stronger restriction on changing forest land to other purposes as this kind of conversion directly affects the quality of the environment. This restriction on forest conversion should come along with stricter penalties and punishment for any forest land violation to eliminate the damage to the existing forest area. Moreover, people who cultivate and protect forest lands should be paid accordingly or rewarded to encourage their willingness to keep forests and reduce environmental degradation.

6.3. The SeABM approach and land use policy making

Another important objective of this research was to develop a tool for modelling land use change in a way that would provide policy makers with the ability to predict land use change under different policy options. The SeABM in this research was able to carry it out with simulations of different policy options. This is an important improvement on what was previously available because it can be quickly deployed with available spatial data with some "easy to collect" demographic measures from households. Policy makers could use this model to compare the different effects of policy options on land use changes and quality of environment. They could have the results visualised in the form of land use maps which would assist them to make better decisions.

Estimating the dynamics of land use, landscape transformations and their influence on environmental quality is complicated and challenging. Since land uses are not spatially uniform and transform continuously over time, they vary according to the position and configuration of land use types and other factors. Both spatial and demographic factors may affect individual behaviours in making land use decisions. In this case, SeABM has shown its ability to handle the interaction among biophysical and socio-economic factors under the hypothesized decision-making behaviour of agents and mechanisms of their feedback to changes. This approach has attempted to solve the difficulty in quantifying the land use change processes over time and space, which could improve the quality of land use policy making. It would save time for planning purposes by allowing comparison of different policy alternatives instantly, rather than using costly and time-consuming surveys. And more importantly, it would reduce the possible unwanted outcomes on environmental quality from any change in policy. The benefit of the model justified the effort put into its construction.

The model can be rebuilt as an integrated tool in ArcGIS, which would be easily deployed by policy makers from provincial level or higher, who may know GIS practice but do not have a modelling background. They may need training to interpret outcomes from the model. Variables for this model are available or can be collected from any area of interest. Policy makers should estimate the scope and the scale of data collection. Policy makers would experience a flexible way to explore the environmental outcomes of different land use decisions across the landscape.

However, policy makers should be aware that, like other modelling, this SeABM has to rely on assumptions. For this research, the model assumed that a household had considered market factors in its decision making. This SeABM also assumed that land tenure or ownership did not affect the decision-making behaviour of households. Although the outcomes of this model would provide rich information and assist the decision-making process, policy makers would still need to consult other sources of information to have an overall picture before making final decisions. Interpretation and judgements of the model's outcomes would rely on the experience and expertise of the policy makers.

6.4. Gaps in the methodology

All models are abstractions of reality, and some work in certain circumstances better than others (Hagedorn et al. 2005). Like many other scientific models, the SeABM in this research tried to represent the phenomena of land use dynamics in a logical and objective way, while focusing on simplifying reflections of reality in order to be useful despite the approximations and assumptions (Box et al. 1987). This SeABM has been developed as a trial to test the possibility of combining spatial and non-spatial factors in modelling the land use decision making process. Since this is the first version applied to this study area, it has found gaps to be explored in further research.

The first gap that could be addressed is the quality of input data for the model. Secondary data including spatial and non-spatial data had been collected from different sources and there was inconsistency across data sets regarding data quality and reliability. It was difficult to incorporate those data to fit into the SeABM and the quality of input data must have affected the econometrical modelling. This may have caused some of the potential independent variables in the regression model to be removed from further analysis. To ensure the quality of the input data, it is suggested they are intentionally collected for the research or there should be consistency among different data from different sources.

ABMs in general, and this SeABM in particular, is well incorporated with simulations, and these incorporations are different from other analytical models. The quality of simulation results in this research depended on how the SeABM had been constructed. There are three aspects in quality control of this SeABM: calibration, verification, and validation. A good discussion in LUCC Report Series No. 6 (Parker et al. 2001) pointed out that most ABM in LUCC research has a clear target of implementing

ABM in a conceptual model that concerns land-use and land-cover change and all of them face a challenge in assuring the quality of modelling.

The difficulty with validating the model is the second gap that needs to be addressed. The only way to know how accurately the model performed is analysing the satellite image in 2020 and comparing the classified land uses with the simulated ones. The classified image must help to calibrate the land use conversion module to reflect the real changes in the study area. What has been used in this research so far was the assumption that the future land use changes may follow the patterns of land use changes from 1995 to 2010. The pattern of land use changes from 2010 to 2020 may be different from what happened in the past. In retrospect, the 2010 image could have been used to validate a similar model starting from 1995, however, there was no demographic data collected in the past to adjust such a model.

The model was subject to structural validation to assess how well the software model represented the conceptual model in terms of programming and outcome validation. It focuses on how well model outputs characterize the target system or the truthfulness of a model with respect to its problem domain. The model was verified by ensuring the proper functioning of its underlying programming or, as pointed out by Parker et al. (2003), verification means ensuring the system was correctly built. Filatova et al. (2013) discussed challenges in verification and validation of ABM. This model always contains stochastic element and when aggregating behaviour from micro scale to macro scale with too many parameters makes the outcomes sensitive. However, if an ABM could be assessed and validated, it would model the real world better.

This research adopted a validation framework to validate model processes and components during model development, which involved a survey at the study site and involvement of stakeholders. During construction of the model, it was verified by sensitivity analysis with different input parameters to test the outputs of the model. Using different scenarios in simulation was also a method to verify the model in simulation mode, to make sure the outputs were mapped according to the parameters of different scenarios. Paul and Bridget (2009) suggested that to have a good validation of ABMs, studies should have a clear description of the phenomena to be explained by the model and test for the simplest possible realistic agent rules of behaviour. When a model is well validated it will provide a strong base to enable that model's comparison and acceptance.

Regarding calibration, the model was calibrated by fitting it to appropriate data before running. But it could be done more accurately if the simulation were restructured to predict land use in five years (from

2010 to 2015) instead of ten years.A classified satellite image of land use in 2015 could then be used to compare the simulated result with real land use, and from that calibrate decision rules and land use change modules in the main simulation framework. However, by the time the analysis was done in 2014 there was no chance to get a 2015 image and since time and budget are limited for this research, this kind of calibration can only be done as an extension to this research.

There are also other challenges that this research needs to consider improving the quality of the SeABM. Scale-related problems have been widely recognized as common challenges for verification and validation with ABMs in LULC. The effect of varying spatial resolution on data analysis have been detected in other studies (Parker et al. 2001). This research tried to maintain 20 metre spatial resolution in data analysis and simulation as a choice after reviewing other research and considering available data.

Another challenge for ABMs lies in abstraction, since many outcomes of human interaction, such as trust or learning, are difficult to implement. Since abstract outcomes are ill-defined and not easily measured, validating them is difficult. This research used the solution discussed by Bousquet et al. (1999) – using expert and stakeholder interviews that provide a sense of how emergent outcomes are related to model structure and processes.

This research faced other challenges in configuring the SeABM so that it can be more easily applied to other areas. The data collection, processing and analysis for this research have shown that it is complicated to use both spatial and non-spatial data to run the model and the simulation. And even if data are available, the quality and consistency of data are still a challenge (as addressed above). Any new area will need a socioeconomic survey at the household level to get the demographic data for the econometric model. Experience from this research has shown that the questionnaire for the survey could be shorter and focus more on demographic information, as land use history can be obtained through satellite images if they are available for the area of interest. It would be easier to apply this model to other cases if all its steps, procedures and data structure could be clearly documented. The flexibility of the model structure, its ease of programming, and the fact that it can be run as a script tool in ArcGIS allows users the ability to tune the decision-making parameters, such as land use change modules, or parameters for each scenario to test the potential outcomes of different scenarios.

Making simplifications and assumptions is essential in simulations as models are expected to capture the general trends or core processes of phenomena rather than trying to duplicate all processes which may have not been well understood or too complicated to express by mathematical formulas or different methodologies. The mechanism of decision making in this research has been much simplified compared to real human behaviours. Households or agents were assumed to make decisions in a hierarchical order. When an agent makes decisions by using the three tiers of decision-making hierarchy, it considered only a few potential factors. In the first tier, demographic data such as cash and labour availability were used to characterise households' further action. Although being simplified, it is assumed that finance and labour are always the first thoughts for local farmers. Agricultural activities are seasonal and labour intensive and moreover, credit accessibility is limited for remote rural areas, which is why households need to be sure about their situation beforehand. The demographic characteristics define the ability of each household to make any LULC changes.

The simple household profiles based on the cash rate and the labour rate in each household could reduce the sensitivity of many scenarios in the simulations, such as the low population growth rate, the high-income growth rate and the financial support scenarios. This research compared cash and labour rates to define the household profiles depending on whether they were negative or positive. This reduced the complexity of the first tier in the decision hierarchy. However, it could also have reduced the possibility of diversifying the household profiles as well as their associated land use conversions.

What makes this SeABM different from other ABMs is that agents did not move in their environment (or study area). They only considered their locations and other spatial related features to make decisions on land use changes. In other words, autonomous agents have control over their actions and internal state to achieve goals without changing their positions. The "exchange land" or "sell/buy land" behaviour was excluded from this SeABM to make it simpler.

An important assumption of land ownership has been made for this research. According to the Land Law 2003, land can be recalled for socio-economic development by the State. The instability of land use rights has been an issue for many developing areas in Vietnam, especially where industrialization and urbanization are promoted. However, for a remote area like My Lam commune, the land tenure seems to be stable, so it was assumed that households did not consider land tenure or ownership while making decisions about land use changes.

6.5. Directions for further research

This research has reassured the ability of the SeABM in land use simulation; the question now is what to do next. There are aspects of the land use simulation and land use decision making behaviour that have not been covered in the present research but are considerably worthwhile to investigate in the near future. Those aspects can be investigated from both the academic viewpoint as well as the technical viewpoint. Challenges and gaps of this research revealed that future patterns, variations and trends of land use need to be carefully considered.

The household profiles which define the land use preferences needs to be studied more deeply. This research has simplified the way household profiles are constructed. The discussion and observation during the survey had assumed that households initially consider the cash balance and the labour balance before drawing any decision. However, a simple threshold of 0 was set to define four profiles. In practice, households having negative rates close to 0 may react like households having positive rates slightly above 0. Future research can work out better thresholds and link them to more accurate household profiles. The refinement of household profiles may improve the sensitivity of some scenarios in simulations.

Land tenure and market signals are two interesting factors that may have an influence on land use decision making. The scope and the budget of this research did not allow study of those factors more deeply. As mentioned in the beginning of this research, private ownership of land is not permitted in Vietnam. The land is owned by the State government and is allocated to households for their settlement and cultivation. Despite many land use rights having been granted for the households, it is likely the long-term lease (50 years for forest land and shorter for agricultural land) with the right to renew the tenure, the state government reserves its right to recall the land for different development purposes and the compensations sometimes do not reflect the willingness of people who must give the land back. The uncertainty in land tenure may affect the way households decide to use their land, forcing them to exploit the land to take as much as they can before giving land back rather than making a long-term commitment. And when households try to make a quick return, the market drivers affect their decisions. Households chase after different crops with an expectation of getting a higher income. However, households often lack the ability to sense the delays and uncertainties from the market which makes the land use decision ineffective. The current research assumed that land tenure and market drive had not been considered by the households when choosing crops and land use types, but those factors can be estimators for the land use change.

Climate change has become an essential topic in Vietnam. It has been integrated in many aspects of land use planning. However, during the survey, households did not express their perceptions of climate change impacts. They gradually tune their decision making and activities towards the short-term weather change rather than thinking far into the future. It would be interesting to explore how climate change adaptation may affect land use decision making behaviour.

Along with the academic side, there are some technical improvements which can be carried out in future research. The scripts can be optimised to bring a better performance for the simulations and after that it will be more helpful if a tool can be built in ArcGIS based on the scripts. The tool will provide a better graphical user interface for other users to try and contribute to the model development. Since the validation for the model in this research needs the future satellite image, future research may have to prepare for that difficulty by finding another alternative to validate and calibrate the model.

In conclusion, by deploying an interdisciplinary approach to solve the research problem, this research faced a formidable set of unique challenges. Households are the crucial component which drives the ABM focus on human actions; this is the major difference from other cellular models used for LULC, which are focused on landscapes and land transitions. And because human behaviours are complicated, this poses a constraint on the SeABM in this research to explain a complex phenomenon. Based on the answers for the three research questions, this research has identified the mechanism of land use decision making and its influencing factors; it successfully set up the SeABM which served as the core to simulate the land use changes in the study area; and it has identified trends of future environmental outcomes, such as changes in forest cover fragmentation, soil erosion and carbon sequestration, under different policy scenarios. It has created the framework where the effects of different policy interventions on environmental outcomes may be tested using simulations. Despite the potential application in land use planning and estimating new development alternatives, there are many gaps to be improved with this research to have better modelling, which better captures the reality of land use decision making behaviour. An improved model will offer both households and planners the opportunity to make more informed decisions.

ACKNOWLEGMENTS

For students, writing a PhD thesis at Lincoln University is a great honour and challenge. I faced a lot of difficulty choosing a research topic which is not only interesting but also practical in Vietnam. This thesis is the joint effort of my supervisors and I, because it would not have the spirit that it has without their invaluable academic and psychological support. I would like to express my special appreciation and thanks to my supervisor Professor Hugh Bigsby and co-supervisor Dr. Crile Doscher; you have been so kind, and such tremendous mentors for me. Your priceless advice on both the research process and my career have encouraged me so much.

The research focuses on links between settlement patterns and land use changes, and environmental degradation. By using the spatially explicit agent-based modelling approach, it helps simulate environmental outcomes under different scenarios of policy alternatives. The research faced several difficulties due to the series of earthquakes in February 2011, which caused many disruptions including periods without power, water and public transport, and periods of closure of Lincoln University. The research was also extended as I kept my full-time job to overcome financial barriers. Without the sympathy of my supervisors, I could never have expected to carry on this research to its conclusion.

Special thanks go to the people in My Lam commune (Cat Tien district, Lam Dong province). They were so friendly and helpful. During my stay in that commune I learnt so much from those people and their experiences, which helped enrich my local knowledge. I also want to express great gratitude to the contacts in Netherlands Development Organisation (SNV) in Vietnam, and Central Division of NIAPP and Centre for Agricultural Policy of Vietnam (CAP) for sharing their data and offering support.

I also want to thank Dr. Bien Nghia Nguyen from the Forest Inventory and Planning Institute of Vietnam (FIPI). His advice and support saved me from losing my way when conducting this research. I would like to express my gratefulness to my parents - Son Nghia Nguyen and Anh Thi Van Hoang, and my sister An Thi Thanh Nguyen for their sentimental support. I give special thanks to my wife Thuy Thi Thanh Nguyen for her understanding and sympathy. And it will never be enough to express my gratefulness to Mr Malcolm King and Mrs Bronwen King in Christchurch, who encouraged me to overcome unexpected difficulties, and who I think of as family.

Gisborne, 18th June 2016

Lan Nghia Nguyen

REFERENCE

ADB (2000). Study on the Policy and Institutional Framework for Forest Resources Management. Rome, Italy and Hanoi, Vietnam: Agriconsulting S.P.A., Asian Development Bank. **TA No.3255 - VIE**.

Adger, W. N. (1998). Observing institutional adaptation to global environmental change: Theory and case study from Vietnam, Working Paper GEC 98-21, The Center for Social and Economic Research on the Global Environment. University of East Anglia and University College London. Norwich and London.

Arthur, W. B., S. N. Durlauf and D. A. Lane (1997). The Economy as an Evolving Complex System II, SFI Studies in the Sciences of Complexity, Addison-Wesley: Reading, MA.

Babbie, E. (2005). The Basics of Social Research. Belmont, CA, Thomson Wadsworth.

Bagni, R., R. Berchi and P. Cariello (2002). "A comparision of simulation models applied to epidemics." Journal of Artificial Societies and Social Simulation **5**(3).

Bhattarai, M. and M. Hammig (2001). "Institutions and the environmental Kuznets Curve for deforestation: A crosscountry analysis for Latin America, Africa and Asia." World Development **29**: 995-1010.

Binh, N. T. (2011). "The trend of Vietnamese household size in recent years." IPEDR **20**: 47-52.

Boserup, E. (1965). The Conditions of Agricultural Growth. London, George Allen & Unwin Ltd.

Boundeth, S., T. Nanseki, S. Takeuchi and T. SATHO (2012). "Land Use Change and Its Determinant Factors in Northern Laos: Spatial and Socio-economic Analysis." 2012 **4**(12).

Bousquet, F., O. Barreteau, C. Le Page, C. Mullon and J. Weber (1999). An environmental modelling approach. The use of multi-agents simulations. (Eds.) Advances in Environmental and Ecological Modelling. F. Blasco and A. Weill. Paris, Elsevier**:** 113-122.

Box, G. E.P. and D. N.R. (1987). Empirical Model-Building and Response Surfaces, Willey.

Brondizio, E., C. Safar and A. Siqueira (2002). "The urban market of Açaí fruit (Euterpe oleracea Mart.) and rural land use change: Ethnographic insights into the role of price and land tenure constraining agricultural choices in the Amazon estuary." Urban Ecosystems **6**(1/2): 67-98.

Brown, D. G., R. Riolo, D. T. Robinson, M. North and W. Rand (2005). "Spatial process and data models: Toward integration of agent-based models and GIS." Journal of Geographical Systems **7**(1): 25-47.

BTNMT (2014). Decision No 1467/QD-BTNMT dated July 21th 2014 of Minister of Natural Resources and Environment. Hanoi, Ministry of Natural Resources and Environment

Butler, R. A. (2006, 27/07/2012). "Threats to rainforest from humankind." Retrieved 02/01/2016, 2016, from http://rainforests.mongabay.com/0803.htm.

Candau, J., S. Rasmussen and K. C. Clarke (2000). A coupled cellular automaton model for land use/land cover dynamics. 4th International Conference on Integrating GIS and Environmental Modeling(GIS/EM4): Problems, Prospects and Research Needs. Alberta, Canada.

Carley, K., N. Altman, E. Casman, D. Fridsma, A. Yahja, L. Chen, B. Kaminsky and D. Nave (2006). "Biowar: Scalable agent-based model of bioattacks." IEEE Trans Syst Man Cybernet **36**(2): 252-265.

Castella, J. C. and D. D. Quang (2002). Doi Moi in the mountains: land use changes and farmers' livelihood strategies in Bac Kan province, Vietnam. Hanoi, Vietnam, Agricultural Publishing House.

Castella, J. C., T. N. Trung and S. Boisau (2005). "Participatory simulation of land-use changes in the northern mountains of Vietnam: the combined use of an agent-based model, a role-playing game, and a geographic information system." Ecology and Society **10**(1): 27.

Castle, C. J. E. and A. T. Crooks (2006). Principles and Concepts of Agent-Based Modelling for Developing Geospatial Simulations Working papers series. London, Center for Advanced Spatial Analysis, University College London.

Chen, Y., S. Khan and Z. Paydar (2009). "To Retire or Expand? A Fuzzy GIS-based Spatial Multi-criteria Evaluation Framework for Irrigated Agriculture." Irrigation and Drainage **59**: 174-188.

Chisholm, M. (1962). Rural Settlement and Land Use: An Essay in Location. London, Hutchison University Library.

Clarke, G. (2001). "From Ethnocide to Ethnodevelopment? Ethnic Minorities and Indigenous People in Southeast Asia." Third World Quaterly **22**(3): 413-436.

Clarke, K. C. (2014). Cellular Automata and Agent-Based Models. Handbook of Regional Science. M. M. Fischer and P. Nijkamp. Berlin, Heidelberg, Springer Berlin Heidelberg**:** 1217-1233.

Clarke, K. C. and L. Gaydos (1998). "Loose coupling a cellular automaton model and GIS: long-term growth prediction for San Francisco and Washington/Baltimore." International Journal of Geographical Information Science **12**(7): 699-714.

Clarke, K. C., S. Hoppen and L. Gaydos (1997). "A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area." Environment and Planning B: Planning and Design **24**: 247-261.

Cliff, A. D., P. Hagget and and R. Martin (1997). "Michael Chisholm: An appreciation." Regional Studies **31**(3): 205-210.

Coe, C. A. (2008). A Tragedy, But No Commons: The Failure of "community-based" Forestry in the Buffer Zone of Tam Dao National Park, Vietnam, and the Role of Household Property Rights and Bureaucratic Conflict. Los Angeles, University of California.

Collier, N. (2000). Repast: An extensible framework for agent simulation. Swarmfest 2000. Logan, Utah.

Cosslett, T. and P. Cosslett (2013). Water Resources and Food Security in the Vietnam Mekong Delta, Springer International Publishing.

Couclelis, H. (1985). "Cellular worlds: A framework for modeling micro-macro dynamic." Environment and Planning A **17**: 585-596.

Cramer, W., A. Bondeau, S. Schaphoff, W. Lucht, B. Smith and S. Sitch (2004). "Tropical forests and the global carbon cycle: impacts of atmospheric carbon dioxide, climate change and rate of deforestation." Philosophical Transactions of the Royal Society B: Biological Sciences **359**(1443): 331-343.

Culas, R. J. (2007). "Deforestation and the environmental Kuznets curve: An institutional perspective." Ecological Economics **61**(2-3): 429-437.

Dang, A., S. Golstein and J. McNally (1997). "Internal Migration and Development in Vietnam." International Migration Review **31**(2): 312-337.

Dang, A., C. Tacoli and X. Hoang (2003). Migration in Vietnam. Regional Conference on Migration, Development and Pro-poor Policy Choice in Asia. Dhaka, Banglades.

De Koninck, R. (1999). Deforestation in Vietnam, International Development Research Center, Ottawa, Canada.

Deadman, P. (2004). Exploring the Role of Communication in an Agent Based Simulation of Land Use Change in the Amazon. The Association of American Geographers Annual Meeting. Philadelphia.

Deadman, P., D. Robinson, E. Moran and E. Brondizio (2004). "Colonist household decision making and land-use change in the Amazon Rainforest: an agent-based simulation." Environment and Planning B: Planning and Design **31**(5): 693-709.

Deadman, P. J., D. T. Robinson, E. Moran and E. Brondizio (2004). "Effects of Colonist Household Structure on Land Use Change in the Amazon Rainforest: An Agent Based Simulation Approach." Environment and Planning B: Planning and Design **31**(5): 693-709.

Duong, N. (2009). The population of Vietnam through periods (Dân số Việt Nam qua các thời kỳ). VnEconomy. Hanoi.

Duy, T., L. Robin, J. Lu and M. James (2010). Population and Natural Resources. Case study: Is population growth responsible for the lost of rainforests? Focus" Southeast Asia, AAG Center for Global Geography Education.

Engelen, G., S. Geertman, P. Smits and C. Wessels (1999). Dynamic GIS and strategic physical planning support:a pratical application, Chapter 5. Geographical Information and Planning. J. C. H. Stillwell, S. Geertman and S. Openshaw. Berlin, Springer**:** 87-111.

Enright, A. (2013). REDD+ compensation packages in Lam Dong Province, Vietnam: Assessing the preferences of forest communities. London, IIED.

Epprecht, M., A. Heinimann, N. Minot, D. Muller and T. Robinson (2007). "From Statistical Data to Spatial Knowledge -- informing decision-making in Vietnam." Information Development **23**: 193.

Epstein, J. M. and R. Axtell (1996). Growing Artificial Societies: Social Science from the Bottom Up. Cambrigde, MA, MIT Press.

ESRI (2013). ArcGIS Desktop: Release 10.1, Redlands, CA: Environmental Systems Research Institute.

Evans, T. P., A. Manire, F. de Castro, E. Brondizio and S. McCracken (2001). "A dynamic model of household decison-making and parcel level land cover change in the eastern Amazon." Ecological Modelling **143**(1-2): 95-113.

FAO (2003). Status and trends in mangrove area extent worldwide. Forest Resources Assessment Working Paper No.63. M. Wilkie and F. S, Food and Agriculture Organisation of the United Nations.

FAO (2015). Global Forest Resources Assessment 2015. Desk reference. Rome, Food and Agriculture Organisation of the United Nations.

Filatova, T., P. H. Verburg, D. C. Parker and C. A. Stannard (2013). "Spatial agent-based models for socio-ecological systems: Challenges and prospects." Environmental Modelling & Software **45**: 1-7.

Forman, R. T. T. (1995). "Land mosaics: The ecology of landscapes and regions." Cambridge University Press. United Kingdom: 652.

FPD. (2013). "Vi phạm luật BV&PTR." Retrieved 26/8/2013, 2013, from http://www.kiemlam.org.vn/Desktop.aspx/List/Hanh-vi-vi-pham-Luat-BV-va-PT-rung/.

Gardner, M. (1970). "The fantastic combinations of john conway's new solitare game "Life"." Scient Amer **223**(120-123).

Gaston, K. J. and J. I. Spicer (2004). Biodiversity: An Introduction, Wiley.

Gibbon, A., V. Evan, L. T. Buana, C. N.T, Q. L.T., T. N.T. and R. McNally (2011). The Feasibility of Carbon Financing for Improved Forest Management at Loc Bac State Operating Company, Lam Dong Province, Vietnam. Hanoi, SNV.

Godoy, R., V. Reyes-García, E. Byron, W. R. Leonard and V. Vadez (2005). "The effect of market economies on the well-being of indigenous peoples and on ther use of renewable natural resources." Annual Review of Anthropology **34**(1): 121-138.

Gustafson, E. J. (1998). "Quantifying Landscape Spatial Pattern: What Is the State of the Art?" Ecosystems **1**(2): 143-156.

Ha, D. T. (2001). Balancing Economic and Environmental Concerns in Uplands of Vietnam: A Continuing Challenge. SANREM CRSP Research Scientific Synthesis Conference. Athens.

Hagedorn, R., J. D. Francisco and T. N. Palmer (2005). "The rationale behind the success of multi-model ensembles in seasonal forecasting - I.Bacis concept." Tellus **57A**: 219-233.

Hai, P. S., N. T. Binh and N. M. Dao (2013). Assessing soil erosion rates for a large catchment in the central highlands of Vietnam using fallout radionuclides, Environmental Research Centre, Nuclear Research Institute, Vietnam Atomic Energy Institute.

Hai, P. S. and T. T. Dung (2014). "Application of fallout radionuclides to estimate soil erosion rates at areas having different farming practices in the Lamdong Region." Vietnam Soil Science **43**.

Haila, Y. (1999). Islands and fragments. Maintaining biodiversity in forest ecosystems. M. L. Hunter. United Kingdom, Cambridge University Press**:** 234-264.

Hamill, L. and N. Gilbert (2015). Heterogeneous demand. Agent-Based Modelling in Economics, John Wiley & Sons, Ltd**:** 29-52.

Harris, L. D. (1984). "The Fragmented Forest: Island Biogeographic Theory and the Preservation of Biotic Diversity." University of Chicago Press, Chicago: 221 pp.

Hill, R., R. Cal and L. Champagne (2006). "Using agent simulation models to examine and investigate search theory against a historical case study." j. Simul **1**(1): 29-38.

Hoan, N. V. (2006). "How to maintain and develop the cultural characteristics of Tay Nguyen ethnics in the globalization background." Scientific Journal of Danang University **14**: 6.

Hoang, M. H., T. H. Do, M. Van Noordwijk, T. T. Pham, M. Palm, X. P. To, D. Doan, T. X. Nguyen and T. V. A. Hoang (2010). An Assessment of Opportunities For Reducing Emissions From All Land Uses – Vietnam preparing for REDD: Final national report. . Nairobi,Kenya., ASB Partnership for the Tropical Forest Margins.

Hoc, D. X. (2002). Drought and It Mitigation Measures (in Vietnamese). Hanoi, Vietnam, Agricultural Publishing House.

Huigen, M. G. A. (2004). "First principles of the MameLuke multi-actor modelling framework for land use change, illustrated with a Philippine case study." Journal of Environmenal Management **72**(1-2): 5-21.

Huynh, T. (2014). Cat Loc Landscape – Cat Tien National Park Pro-Poor REDD+ Project, Vietnam. REDD+ on the ground. S. S. A. Erin O Sills, Claudio de Sassi, Amy E Duchelle, Demetrius L Kweka, Ida Aju Pradnja Resosudarmo and William D Sunderlin. Indonesia, CIFOR.

Irwin, E. G. and J. Geoghegan (2001). "Theory, data, methods: developing spatially explicit economic models of land use change." Agriculture, Ecosystems & Environment **85**(1-3): 7-24.

IUCN (2006). IUCN Red List of Treatened Species, IUCN.

Jamieson, N., T. C. Le and A. Rambo (1998). The development crisis in Vietnam's mountains. Special Report no.6. Honolulu, Hawaii, USA, East-West Center.

Janssen, M. A. and E. Ostrom (2006). "Empirically-based, agent-based models." Ecology and Society **11**(2): 37.

Jepma, C. J. (2014). Tropical Deforestation: A Socio-economic Approach, Taylor & Francis.

Kaimowitz, D. and A. Angelsen (1998). Economic models of tropical deforestation. A review. Bogor, Indonesia, Center for International Forestry Research.

Kocabas, V. and S. Dragicevic (2006). "Assessing cellular automata model behaviour using a sensitivity analysis approach." Computers, Environment and Urban Systems **30**(6): 921-953.

Koyunen, C. and R. Yilmaz (2009). "The Impact of Corruption on Deforestation: a Cross-Country Evidence." The Journal of Developing Ideas.

Lamdong (2010). Report on Environmental Condition of Lamdong province. Dalat.

Lamdong. (2012). "Xa My Lam."   Retrieved 01/04/2012, 2012, from http://www.lamdong.gov.vn/vi-VN/a/cattien/ubnd-huyen/xa-thitran/Pages/xa-MyLam.aspx.

Lamdong. (2015). "Quản lý, bảo vệ rừng: Yếu kém."   Retrieved 12/12/2015, 2015, from http://baolamdong.vn/phapluat/201511/quan-ly-bao-ve-rung-yeu-kem-2643237/.

Lang, C. (2001). Deforestation in Vietnam, Laos and Cambodia. Deforestation, Environment, and Sustainable Development: A Comparative Analysis. D. K. Vajpeyi. Westport, Connecticut and London, Praeger**:** 111-137.

Laurance, W. F. (2004). "Forest–climate interactions in fragmented tropical landscapes." Phil. Trans. R. Soc. Lond. **B 359**: 345-352.

Le, Q. B. (2005). Multi-Agent System for Simulation of Land-use and Land-cover Change: A Theoretical Framework and Its First Implementation for An Upland Watershed in the Central Coast of Vietnam. Ecological and Development Series. Göttingen, Cuvillier Verlag Göttingen. **29**.

Le, Q. B. and S. Park (2008). "Land-Use Dynamic Simulator (LUDAS): A multi-agent system model for simulating spatio-temporal dynamics of coupled human–landscape system. I. Structure and theoretical specification." Ecological Informatics **3**.

Levy, P. E., M. G. R. Cannell and A. D. Friend (2004). "Modelling the impact of future changes in climate, CO2 concentration and land use on natural ecosystems and the terrestrial carbon sink." Global Environmental Change **14**(1): 21-30.

Liu, Y. (2008). Modelling Urban Development with Geographical Information Systems and Cellular Automata, CRC Press.

Loi, N. K., N. T. T. Thuy, N. T. Huyen and V. M. Tuan (2010). Integration of GIS and AHP Techniques for Analyzing Land Use Suitability in Di Linh District, Upstream Dong Nai Watershed, Vietnam. Agriculture & Development Discussion Paper Series, SEARCA.

Long, J. S. (1997). Regression Models for categorical and limited dependent variables. Thousand Oaks, CA, Sage Publications.

Long, J. S. and J. Freese (2005). Regression Models for Categorial Dependent Variables Using Stata. Second Edition.

Loyn, R. H. and C. McAlpine (2001). Spatial patterns and fragmentation: indicators for conserving biodiversity in forest landscapes. Criteria and indicators for sustainable forest management. R. J. Raison, A. G. Brown and D. W. Flinn. United Kingdom, CABI Publishing**:** 391-442.

LPSO (2005). Lamdong Statistical Book 2005: Administrative unit, Area and Population, Lamdong Statistical Office.

LPSO (2006). Lamdong Statistical Book 2006: Administrative unit, Area and Population, Lamdong Statistical Office.

LPSO (2007). Lamdong Statistical Book 2007: Administrative unit, Area and Population, Lamdong Statistical Office.

LPSO (2008). Lamdong Statistical Book 2008: Population and Employment - Area, population and population density in 2008 by district, Lamdong Statistical Office.

LPSO (2009). Lamdong Statistical Book 2009: Population and Employment - Area, population and population density in 2009 by district, Lamdong Statistical Office.

Macal, C. M. (2004a). Emergent structures from trust relationships in supply chains. Proceedings of agent 2004: Conference on Social Dynamic: Interaction, Reflexivity and Emergence. C. M. Macal, D. Sallach and M. North. Chicago, IL, 7-9 October, Argonne National Laboratory**:** 743-760.

Macal, C. M. and M. J. North (2010). "Tutorial on agent-based modelling and simulation." Journal of Simulation **4**: 151-162.

Macal, M. C. and J. M. North (2010). "Tutorial on agent-based modelling and simulation." Journal of Simulation **4**(3): 151-162.

Maphill (2011).

MARD (2001). Synthesis report: five million hectare reforestation program partnership. Hanoi, Ministry of Agriculture and Rural Development, International Co-operation Department, February.

Marsh, W. E. and R. Hill (2008). "An initial agent behavior modeling and definition methodology as applied to unmanned aerial vehicle simulation." Int J Simul Process Model **4**(2): 119-129.

Martine, N. (2012). "GIS and Multi-Criteria Decision Analysis for Land Use Resource Planning." Journal of Geographic Information System **4**: 341-348.

Matthews, R. B., N. G. Gilbert, A. Roach, J. G. Polhill and N. M. Gotts (2007). "Agent-based land-use models: a review of applications." Landscape Ecology **22**(10): 1447-1459.

Mayle, F. E., D. J. Beerling, W. D. Gosling and M. B. Bush (2004). "Responses of Amazonian ecosystems to climatic and atmospheric carbon dioxide changes since the last glacial maximum." Phil. Trans. R. Soc. Lond **B 359** 499–514.

McCracken, S., E. Brondizio, D. Nelson, E. Moran, A. Siqueira and Rodriguez-Pedraza (1999). "Remote sensing and GIS at farm property level: Demography and deforestration in the Brazilian Amazon." Photogrammetric Engineering and Remote Sensing **65**: 1311-1320.

McCracken, S., A. Siqueira, E. Moran and E. Brondizio (2002). Land use patterns on an agricultural frontier in Brazil: insights and examples from a demographic perspective. Deforestation and Land Use in the Amazon. C. Wood and R. Porro. Gainesville, University Press of Florida**:** 162-196.

McGarigal, K. and B. Marks (1993). FRAGSTAT: spatial pattern analysis program for quantifying landscape structure (software), Dept. Forest Science, Oregon state University.

McGarigal, K. and B. J. Marks (1995). FRAGSTATS - Spatial pattern analysis program for quantifying landscape structure. Version 2.0. Oregon, Forest Science Department, Oregon State University.

McGarigal, K., S. Tagil and S. A. Cushman (2009). "Surface metrics: an alternative to patch metrics for the quantification of landscape structure." Landscape Ecol **24**: 443-450.

Mena, C. F. (2008). "Trajectories of Land-use and Land-cover in the Northern Ecuadorian Amazon: Temporal Composition, Spatial Configuration, and Probability of Change." Photogrametric Engineering & Remote Sensing **74**(6): 14.

Ménard, A. and C. M. Macal (2005). "Exploration of spatial scale sensitivity in geographic cellular automata." Environment and Planning B: Planning and Design **32**: 693-714.

Meyfroidt, P. and E. Lambin (2008a). "Forest transition in Vietnam and its environmental impacts." Global Change Biology **14**(6): 1319-1336.

Millington, J., R. u. l. Romero-Calcerrada, J. Wainwright and G. Perry (2008). "An Agent-based Model of Mediterranean Agricultural Land-use/Cover Change for Examining Wildfire Risk." Journal of Artificial Societies and Social Simulation **11**(4): 4.

Minar, N., R. Burkhart, C. Langton and M. Askenazi (1996). The SWARM simulation system: a toolkit for building multi-agent simulations. Working Paper 96-06-042. Santa Fe Institute.

Moffat, J., J. Smith and S. Witty (2006). "Emergent behaviour: theory and experimentation using the MANA model." J. Appl Math Decis Sci **10**: 1-13.

Moreno, N., A. Ménard and D. J. Marceau (2008). "VecGCA: A vector-based geographic cellular automata model allowing geometric transformations of polygons." Environment and Planning B **35**: 647-665.

Muller, D. (2004). From agricultural expansion to intensification: Rural development and determinants of land-use change in the Central Highlands of Vietnam. Tropical Ecology Support Program (TOEB). G. Hoebart. Eschborn, GTZ.

Müller, D. and D. K. Munroe (2005). "Tradeoff between Rural Development Policies and Forest Protection: Spatially Explicit Modelling in the Central Higlands of Vietnam." Land Economics **81**(3).

Muller, D. and M. Zeller (2002). "Land use dynamics in the central highlands of Vietnam: a spatial model combining village survey data with satellite imagery interpretation." Agricultural Economics **27**: 333-354.

Mulvaney, D. (2011). Green Politics: An A-to-Z Guide, SAGE Publications.

Nagendra, H. and J. Southworth (2009). Reforesting Landscapes: Linking Pattern and Process, Springer Netherlands.

Nam, T. (2008). Market and Quality Assessment of Pepper in Vietnam, Sustainable Management of Natural Resources in Central Vietnam.

Newbery, D. M., T. H. Clutton-Brock and G. T. Prance (2000). Changes and Disturbance in Tropical Rainforest in South-East Asia, Royal Society.

North, M. and C. M. Macal (2007). Managing Business Complexity: Discovering Strategic Solutions with Agent-Based Modeling and Simulation. Oxford, UK, Oxford University Press.

O'Sullivan, D. (2001). "Exploring spatial process dynamics using irregular cellular automaton models." Geographical Analysis **33**(1): 1-18.

O'Sullivan, D. and P. M. Torrens (2000). Cellular models of urban systems. In Theoretical and Practical Issues on Cellular Automata. Fourth International Conference on Cellular Automata for Research and Industry, Karlsruhe, Springer-Verlag.

OECD (2015). OECD Food and Agricultural Reviews Agricultural Policies in Viet Nam 2015, OECD Publishing.

Ogonowski, M. and A. Enright (2013). Cost implications for pro-poor REDD+ in Lam Dong Province, Vietnam: opportunity costs and benefit distribution systems. London, IIED.

Oscar, S. (2003). The ethnography of Vietnam's Central Highlanders: A historical contextualization 1850 - 1900. Honolulu, University of Hawai's Press.

Park, H. M. (2009). Regression Models for Binary Dependent Variables Using Stata, SAS, R, LIMDEP, and SPSS, The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing, Indiana University."

Parker, C. D., T. Berger and M. S. Manson (2001). Agent-Based Models of Land-Use and Land-Cover Change. LUCC Report Series No. 6. Belgium, LUCC International Project Office 2002.

Parker, D. C., T. Berger and S. M. Manson (2002). Meeting the Challenge of Complexity: Proceedings of the Special Workshop on Agent-Based Models of Land-Use/Land-Cover Change. Santa Barbara: CIPEC/CSISS Publication CCR-3. http://www.csiss.org/maslucc/ABM-LUCC.htm.

Parker, D. C., S. M. Manson, M. A. Janssen, M. J. Hoffmann and P. J. Deadman (2003). "Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review." Annals of the Association of American Geographers **93**(2): 314-337.

Patz, J. A., P. Daszak, G. M. Tabor, A. A. Aguirre, M. Pearl, J. Epstein, N. D. Wolfe, A. M. Kilpatrick, J. Foufopoulos, D. Molyneux, D. J. Bradley and E. Members of the Working Group on Land Use Change Disease (2004). "Unhealthy Landscapes: Policy Recommendations on Land Use Change and Infectious Disease Emergence." Environmental Health Perspectives **112**(10): 1092-1098.

Paul, O. and R. Bridget (2009). Validation and Verification of Agent-Based Models in the Social Sciences. Epistemological Aspects of Computer Simulation in the Social Sciences

F. Squazzoni. Berlin, Springer Berlin Heidelberg. **5466:** 130-140.

Peduzzi, P., J. Concato, E. Kemper, T. R. Holford and A. R. Feinstein (1996). "A simulation study of the number of events per variable in logistic regression analysis." Journal of Clinical Epidemiology **49**: 1373-1379.

Pehu, E. (1998). Upland Agriculture - Regional Report. Regional Environmental Technical Assistance 5771, Poverty Reduction & Environmental Management in Remote Greater Mekong Subregion Watersheds Project (Phase I).

Pender, J., P. Jagger, E. Nkonya and D. Sserunkuuma (2001). Development pathways and land management in Uganda: Causes and implications. Environment and Production Technology Division. Discussion paper No 85, International Food Policy Research Institute**:** 1-88.

Prime Minister (2005). Poverty line applied for period of 2006-2010. (Quyết định về việc ban hành chuẩn nghèo áp dụng cho giai đoạn 2006-2010). 170/2005/QĐ-TTg. Government of Social Republic of Vietnam. Hanoi.

Prime Minister (2008). Trial for payment for forest environmental services (Thí điểm chi trả dịch vụ môi trường rừng). 380/QĐ-TTg. Government of Social Republic of Vietnam. Hanoi.

Rempell, R. S., D. Kaukinen and A. P. Carr (2012). Patch Analyst and Patch Grid. Thunder Bay, Ontario, Ontario Ministry of Natural Resources. Centre for Northern Forest Ecosystem Research.

Rochelle, J. A., L. A. Lehmann and J. Wisniewski (1999). "Forest fragmentation: Wildlife and management implications." Brill Academic Publishers, Leiden, The Netherlands.

Rounsevell, M., D. Robinson, D. Murray-Rust, V. Rieser and V. Milicic (2010). Conceptual agent-based model of RUR land-use dynamics. Koper case study. Peri-urban land use relationships - strategies and sustainability assessment tools for urban-rural linkages, integrated project, Contract No, 036921. S. I. Assessment.

Sala, O. E., F. Stuart Chapin , III, J. J. Armesto, E. Berlow, J. Bloomfield, R. Dirzo, E. Huber-Sanwald, L. F. Huenneke, R. B. Jackson, A. Kinzig, R. Leemans, D. M. Lodge, H. A. Mooney, M. n. Oesterheld, N. L. Poff, M. T. Sykes, B. H. Walker, M. Walker and D. H. Wall (2000). "Global Biodiversity Scenarios for the Year 2100." Science **287**(5459): 1770-1774.

Samat, N. (2007). "Integrating GIS and Cellular Automata spatial model in evaluating urban growth: Prospects and challenges." Jurnal Alam Bina, Jilid 09 **1**.

Saunders, D. A., R. J. Hobbs and C. R. Margules (1991). "Biological consequences of ecosystem fragmentation: A review." Conserv. Biol. **5**: 18-32.

Schaeffer, R. K. (2003). Understanding Globalization: The Social Consequences of Political, Economic, and Environmental Change, Rowman & Littlefield Publishers.

Schmidheiny, K. and U. Basel (2014). Binary Response Models.

SNV (2010). Deforestation Drivers and Community Assessment: Tien Hoang and Dong Nai Thuong communes, SNV.

Soares-Filho, B., G. Cerqueira and C. Pennachin (2002). "dinamica—a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier." Ecological Modelling **154**(3): 217-235.

StataCorp (2009). Stata Statistical Software: Release 11, College Station, TX: StataCorp LP.

Sunderlin, W. D. and T. Huynh (2005). Poverty alleviation and forests in Vietnam, CIFOR.

Takenori, T., M. Asako and F. H. Shigeaki (2010). "Derivation of a yearly transition probability matrix for land-use dynamics and its application." Landscape Ecology **25**(4): 561-572.

Takeyama, M. and H. Couclelis (1997). "Map dynamics: integrating cellular automata and GIS through Geo-Algebra." International Journal of Geo-Information Science **11**: 73-91.

Templeton, S. and S. J. Scherr (1997). Population pressure and microeconomy of land management in hills and mountains of developing countries. EPTD discussion papers with No 26. Washington DC, International Food Policy Research Institute.

Terborgh, J. (1989). "Where Have all the Birds Gone?" Princeton University Press, New Jersey: 207pp.

Thiennhien.net. (2010). "Forests in the Central Highlands have been being destroyed. (Rừng Tây Nguyên vẫn bị tàn phá)." Retrieved 17/08/2011, from http://www.thiennhien.net/2010/06/25/rung-tay-nguyen-van-bi-tan-pha/.

Thuy, N. T. B., Thanh and Trung (2011). Influence of PES to socioeconomic in Lamdong province, The 3rd Southeast Asia Workshop on Payments for Ecosystem Services, Banda, Indonesia 12-14th June 2011.

Tinh, D. N. (2006). Coping with drought in the central highlands - Vietnam. PhD, Technical University of Denmark.

Tobias, R. and C. Hofmann (2004). "Evaluation of free Java-libraries for social-scientific agent based simulation." Journal of Artificial Societies and Social Simulation **7**(1).

UCLA. (2014). "Stata Annotated Output Probit Regression." Retrieved 12 September 2014, 2014, from http://www.ats.ucla.edu/stat/stata/output/Stata_Probit.htm.

UN-REDD (2013). Institution and Context Analysis to inform the PGA for REDD+ in Viet Nam.

UN-REDD. (2015). "About REDD+." Retrieved 20/11/2015, 2015, from http://www.un-redd.org/aboutredd.

Vajpeyi, D. K. (2001). Deforestation, Environment, and Sustainable Development: A Comparative Analysis, Praeger.

van der Werf, G. R., D. C. Morton, R. S. DeFries, J. G. J. Olivier, P. S. Kasibhatla, R. B. Jackson, G. J. Collatz and J. T. Randerson (2009). "CO_2 emissions from forest loss." Nature Geoscience **2**(11): 737-738.

Van Kooten, G. C. and S. Wang (2003). Institutional, Social and Economic factors behind deforestation: A cross-country examination. XII World Forestry Congress. Quebec city, Canada.

Verburg, P. H., W. Soepboer, R. Limpiada, M. V. O. Espaldon, M. A. Sharifa and A. Veldkamp (2002). "Modelling the spatial dynamics of regional land use: The CLUE-s model." Environmental Management **30**: 391-405.

VN Embassy. (2011). "Learn about Vietnam: Culture - Ethnic groups." Retrieved Aug 10, 2011, from http://www.vietnamembassy-usa.org/learn_about_vietnam/culture/ethnic_groups/.

Wade, T. G., K. H. Riitters, J. D. Wickham and K. B. Jones (2003). "Distribution and causes of global forest fragmentation." Conservation Ecology **7**(2).

Wainwright, J. (2008). "Can modelling enable us to understand the role of humans in landscape evolution?" Geoforum **39**(2): 659-674.

White, R., G. Engelen and I. Uljee (1997). "The use of constrained cellular automata for high-resolution modelling of urban land use dynamics." Environment and Planning B **24**: 323-343.

Wilson, E. O. (2003). The Future of Life, Vintage Books.

WorldBank (2005). Vietnam Environmental Monitor. Biodiversity. Washington DC, World Bank.

Wu, F. (2002). "Calibration of stochastic cellular automata: the application to rural-urban land conversions." International Journal of Geographical Information Science **16**(8): 795 - 818.

Yongjin, J., J. Chulmin and P. Soohong (2010). Design of a Dynamic Land-Use Change Probability Model Using Spatio-Temporal Transition Matrix. Lecture Notes in Computer Science. D. T. e. al. **Volume 6010:** 105-115.

Yu, J., Y. Chen and J. P. Wu (2009). Cellular automata and GIS based landuse suitability simulation for irrigated agriculture. 18th World IMACS/MODSIM Congress. Cairn, Australia.

Zang, H. X., P. Mick Kelly, C. Locke, A. Winkels and W. N. Adger (2001). Structure and implication of migration in a transitional economy: Beyond the planned and spontaneous dichotomy in Vietnam. CSERGE Working Paper GEC 01-01.

APPENDIX 1

Questionnaire

Lincoln University, Commerce Falculty

Project: Land use changes and potential impacts on environmental outcomes in Lam Dong province

ID : ....

1. Demography of household:	2. Income from and expenditures for non-farm activities of a household (in VND):
1.1 Where does household originate from? 1- New migrant 2-Local indigenous 1.2 Number of people in household? .... 1.3 What is the age of your oldest child? .... 1.3 What is the highest education level of household head? 1-Lower high school 2-High school 3-Vocational training 4-Higher education 1.4 How many members of household are involved in agricultural activities?	2.1. Estimating annual income of household from non-farm activities? ....VND 2.2. Estimating annual expenditure for household for excluding farming activities? .... VND

3. The land use decision making of household:

3.1 If any of the factors listed below that affected your land use decision during three periods from 1995-2012? How ?

	Factors	Period 1995-2000		Period 2000-2005		Period 2005-2012	
3.1.1	Finance of household	Y		Y		Y	

	Factors	Period 1995-2000		Period 2000-2005		Period 2005-2012	
3.1.2	Labour availability of household	Y		Y		Y	
3.1.3	Price of commodities	Y		Y		Y	
3.1.4	Access to road/water/residential	Y		Y		Y	
3.1.5	Influence from neighbourhood	Y		Y		Y	

3.2 Do you make the decision on land uses changes (1) by yourself or (2) together with other members in household or (3) with other people outside household ?

3.3 How important are those factors on your land use decision?

	Factors	1-not important 5-very important
3.3.1	Finance of household	1 – 2 – 3 – 4 – 5
3.3.2	Labour availability of household	1 – 2 – 3 – 4 – 5
3.3.3	Price of commodities	1 – 2 – 3 – 4 – 5
3.3.4	Access to road/water/residential	1 – 2 – 3 – 4 – 5
3.3.5	Influence from neighbourhood	1– 2 – 3 – 4 – 5

3.4 Please identify your land plots on the commune map and provided some information: (additional pages can be used if household has more than 5 plots)

Plot	Tenure	Area 000m2	Land use in 2012	Labour day per year	Expenditure on inputs per year VND	Estimated output in VND per year	Land use in 1995	Land use in 2000	Land use in 2005

Plot	Tenure	Area 000m2	Land use in 2012	Labour day per year	Expenditure on inputs per year VND	Estimated output in VND per year	Land use in 1995	Land use in 2000	Land use in 2005

4. What land use would you choose if you had the current land use as shown in the pictures?

Picture of current land use	Intended land use	Why and when would you want to change the current land use to intended one
1.Poor forest		
2.Young forest		
3.Medium forest		
4.Mixed forest		
5.Bamboo forest		
6.Rice field		
7. Annual crops (maize, cassava...)		
8. Perennial crops (coffee, pepper...)		

5. Assessing the weights for probability calculation:

If you have to spend “10 effort scale” to make a final decision on land use, how will you consider the below factors?

Factors	Weight
Finance of household	
Labour availability of household	
Price of commodities	
Access to road/water/residential	
Influence from neighbourhood	
Slope	
Elevation	
Sum	**10**

6. Do you think it is necessary to convert some forest areas to other purposes to get a higher financial benefit?
……
……
7. What are the difficulties of converting forests to other purposes when it is necessary? And what would you do to overcome those difficulties?
……
……
……
……
8. Do you think a higher forest protection fee will encourage you, your family and your neighbours to keep the forests? And how much is acceptable?
……
……
……

9. Some basic estimations for spatial parameters:

Access to road/water/residential	“Far” is equal or greater than.................km	“Close” is equal or less than...............km
Slope	“Steeply” is equal or greater than...............%	“Flat” is equal or less than......................%
Elevation	“High” is equal or greater than....................m	“Low” is equal or less than......................m

Thank you for your cooperation!

APPENDIX 2

	
Poor forest	Young forest

	
Medium forest	Mix forest
	
Bamboo forest	Rice field

	
Cassava (annual crop)	Maize (annual crop)
	
Pepper (perennial crop)	Coffee (perennial crop)

	
Cashew (perennial crop)	

APPENDIX 3

```
# SIMULATION OF LANDUSE IN MY LAM COMMUNE, LAM DONG PROVINCE

#Author: Lan Nghia Nguyen
#Lincoln University
#Last update: 22/02/2015

#Parameter for command line:
#                              0.No policy is applied:      "E:/Lan/landuse2010.shp" "0" "0" "0" "0" "0" "0"
Baseline scenario
#                              1.Low population growthrate  "E:/Lan/landuse2010.shp" "1" "0" "0" "0" "0" "0"
#                              2.High income promotion      "E:/Lan/landuse2010.shp" "0" "1" "0" "0" "0" "0"
#                              3.Financial support          "E:/Lan/landuse2010.shp" "0" "0" "1" "0" "0" "0"
#                              4.Cashew promotion           "E:/Lan/landuse2010.shp" "0" "0" "0" "1" "0" "0"
#                              5.Acacia hybrid promotion    "E:/Lan/landuse2010.shp" "0" "0" "0" "0" "1" "0"
#                              6.Payment for forest         "E:/Lan/landuse2010.shp" "0" "0" "0" "0" "0" "1"
#                                4.5.6. can not be activated at the same time

#******************************

# PART I: Import all necessary modules
import arcpy
import random                                 #To create random number
import operator                               #To work with MOD()
import os                                     #To work with folder/file
import ntpath                                 #To work with file/folders
import fnmatch                                #To work with file/folders: function match
import shutil                                 #To work with shutil.copy when backing up org file
import time                                   #To measure processing time
from arcpy import env                         #To set environment
from scipy.stats import norm                  #To calculate the function of cumulative distribution of a normal
distribution, then compare with the probability of being changed from each cell
arcpy.env.overwriteOutput = True

# Set environment settings and inputs for modelling
env.workspace = "E:/LAN"
inpath = arcpy.GetParameterAsText(0)          #The shapefile, can be input through arcpy.GetParameterAsText().
Filename format is "landuseYYYY.???"
    #Policy intervention design:
```

```
lowgrowth = arcpy.GetParameterAsText(1)    # 1=Yes: assigning growthrate = 1.5% and run simulation or 0=No: Keep
current growth rate for baseline is 2.1% and run
highincome = arcpy.GetParameterAsText(2)   # 1=Yes, assign the income local area tries to target every year,
0=No, keep the same level. may need additional variable of "incomegrowth"
credit = arcpy.GetParameterAsText(3)       # 1=Yes poor HH can borrow money to ease the deficit financial
balance. 0=No, no money injected. Goes with "borrow"
procashew = arcpy.GetParameterAsText(4)    # 1=Yes activate land conversion modules which promote cashew
development, 0=No
proacacia = arcpy.GetParameterAsText(5)    # 1=Yes activate land conversion modules which promote acacia hybrid
plantation, 0=No, run basic modules
propfes = arcpy.GetParameterAsText(6)      # 1=Yes activate land conversion modules that promote forest land to
get payment. 0 = No, run basic modules

if procashew == proacacia == propfes == 1 or procashew == proacacia==1 or procashew == propfes == 1 or proacacia
== propfes == 1 :
    print "Error ! Only 1 in 3 promotion programs can not be activated at a time"



#Starting timing
starttime = time.time()

#-------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------
#Land conversion modules (md) definition or function definition: argument of function is LU

def conversion(con1,con2,con3,prf,lu,elv,slp,diswtr,distrans,time,prb,iniprb): #Feed the scenarios to function
con1 = procashew, con2 = procacia, con3 = propfes, prf= profile,lu= row.CLU,elv= row.elv,slp= row.slp,diswtr=
row.diswtr,distrans= row.distrans,time= t
    if con1 == 1 and con2 == 0 and con3 == 0 and prb >= iniprb:         #Activate cashew promotion conversion
        if operator.mod(time,5) == 0 and time < 10:
            if lu == 2:
                if  prf in (3,4) and 170 <= elv <= 242 and 21 <= slp <= 51 and 0 <= diswtr <= 251 and 994 <=
distrans <= 2843:  #md1
                    lu = 3
                else:
                    if prf in (3,4) and 151 <= elv <= 217 and 6 <= slp <= 50 and 10 <= diswtr <= 293 and 948 <=
distrans <= 2992: #md2
                        lu = 5
                    else:
```

```
                        if prf in (1,2,3,4) and 23 <= elv <= 44 and 35 <= slp <= 70 and 0 <= diswtr <= 10 and 
3153 <= distrans <= 3247:#md3
                            lu = 17
            else:
                if lu == 3:
                    if prf == 2 and 75 <= elv <= 321 and 0 <= slp <= 75 and 10 <= diswtr <= 503 and 0 <= 
distrans <= 979: #md456: parameters found acceptable for perenial trees
                        lu = 101 #procashew
                else:
                    if lu == 5 and prf == 2 and 141 <= elv <= 261 and 10 <= slp <= 39 and 0 <= diswtr <= 423 and 
5 <= distrans <= 2973: #md7
                        lu = 101 #procashew
                    else:
                        if lu == 6 and prf == 1 and 128 <= elv <= 191 and 2 <= slp <= 34 and 36 <= diswtr <= 322 
and 14 <= distrans <= 555: #md8
                            lu = 101 #to cashew
                        else:
                            if lu == 9 and prf == 3 and 133 <= elv <= 134 and 2 <= slp <= 3 and 45 <= diswtr <= 
51 and 35 <= distrans <= 53: #md9
                                lu = 112 #to cassava
                            else:
                                if lu == 102: #Since this is procashew, only acacia could be back to bamboo
                                    if prf == 4 and 132 <= elv <= 144 and 1 <= slp <= 31 and 80 <= diswtr <= 349 
and 79 <= distrans <= 205: #md10
                                        lu = 6
                                    else:
                                        if prf == 4 and 132 <= elv <= 132 and 1 <= slp <= 2 and 60 <= diswtr <= 
70 and 88 <= distrans <= 99: #md11
                                            lu = 112 #to cassava
                                else:
                                    if lu in (111,112) and prf == 4 and 132 <= elv <= 134 and 2 <= slp <= 3 and 
59 <= diswtr <= 91 and 34 <= distrans <= 81: #md12
                                        lu = 6
                                    else:
                                        if lu == 17 and prf == 4 and 31 <= elv <= 130 and 50 <= slp <= 74 and 0 
<= diswtr <= 14 and 2562 <= distrans <= 3234: #md13
                                            lu = 6
        else:
            if operator.mod(time,10) == 0:
                if lu == 2:
```

```
                    if prf == 3 and 150 <= elv <= 235 and 15 <= slp <= 49 and 10 <= diswtr <= 36 and 2553 <=
distrans <= 2822:#md14
                        lu = 3
                    else:
                        if prf == 3 and 23 <= elv <= 166 and 12 <= slp <= 64 and 10 <= diswtr <= 20 and 2978 <=
distrans <= 3252: #md15
                            lu = 5
                        else:
                            if prf == 3 and 210 <= elv <= 242 and 27 <= slp <= 41 and 198 <= diswtr <= 290 and
954 <= distrans <= 1038: #md16
                                lu = 6
                else:
                    if lu == 5 and prf == 3 and 105 <= elv <= 300 and 6 <= slp <= 73 and 0 <= diswtr <= 681 and
15 <= distrans <= 3118:#md17
                        lu = 101 #Procashew
                    else:
                        if lu == 6:
                            if prf == 1 and 20 <= elv <= 138 and 0 <= slp <= 59 and 36 <= diswtr <= 282 and 25
<= distrans <= 501: #md18
                                lu = 101 #procashew
                            else:
                                if prf == 1 and 75 <= elv <= 321 and 1 <= slp <= 75 and 20 <= diswtr <= 469 and
5 <= distrans <= 979: #md19
                                    lu = 101 #to cashew
                                else:
                                    if prf == 3 and 117 <= elv <= 280 and 1 <= slp <= 66 and 0 <= diswtr <= 359
and 54 <= distrans <= 3220: #md20
                                        lu = 112 #to cassava
    else:
        if con1 == 0 and con2 == 1 and con3 ==0 and prb >= iniprb:
            if operator.mod(time,5) == 0 and time < 10:
                if lu == 2:
                    if  prf in (3,4) and 170 <= elv <= 242 and 21 <= slp <= 51 and 0 <= diswtr <= 251 and 994 <=
distrans <= 2843:  #md1
                        lu = 3
                    else:
                        if prf in (3,4) and 151 <= elv <= 217 and 6 <= slp <= 50 and 10 <= diswtr <= 293 and 948
<= distrans <= 2992: #md2
                            lu = 5
                        else:
```

```
                            if prf in (1,2,3,4) and 23 <= elv <= 44 and 35 <= slp <= 70 and 0 <= diswtr <= 10
and 3153 <= distrans <= 3247:#md3
                                lu = 17
            else:
                if lu == 3:
                    if prf == 2 and 75 <= elv <= 321 and 0 <= slp <= 75 and 10 <= diswtr <= 503 and 0 <=
distrans <= 979: #md456: parameters found acceptable for perenial trees
                        lu = 102 #proacacia
                else:
                    if lu == 5 and prf == 2 and 141 <= elv <= 261 and 10 <= slp <= 39 and 0 <= diswtr <= 423
and 5 <= distrans <= 2973: #md7
                        lu = 102 #proacacia
                    else:
                        if lu == 6 and prf == 1 and 128 <= elv <= 191 and 2 <= slp <= 34 and 36 <= diswtr <=
322 and 14 <= distrans <= 555: #md8
                            lu = 102 #pro acacia
                        else:
                            if lu == 9 and prf == 3 and 133 <= elv <= 134 and 2 <= slp <= 3 and 45 <= diswtr
<= 51 and 35 <= distrans <= 53: #md9
                                lu = 112 #to cassava
                            else:
                                if lu == 101: #since this is proacacia, only cashew could be converted
                                    if prf == 4 and 132 <= elv <= 144 and 1 <= slp <= 31 and 80 <= diswtr <=
349 and 79 <= distrans <= 205: #md10
                                        lu = 6
                                    else:
                                        if prf == 4 and 132 <= elv <= 132 and 1 <= slp <= 2 and 60 <= diswtr
<= 70 and 88 <= distrans <= 99: #md11
                                            lu = 112 #to cassava
                                else:
                                    if lu in (111,112) and prf == 4 and 132 <= elv <= 134 and 2 <= slp <= 3
and 59 <= diswtr <= 91 and 34 <= distrans <= 81: #md12
                                        lu = 6
                                    else:
                                        if lu == 17 and prf == 4 and 31 <= elv <= 130 and 50 <= slp <= 74
and 0 <= diswtr <= 14 and 2562 <= distrans <= 3234: #md13
                                            lu = 6
        else:
            if operator.mod(time,10) == 0:
                if lu == 2:
```

```
                        if prf == 3 and 150 <= elv <= 235 and 15 <= slp <= 49 and 10 <= diswtr <= 36 and 2553 <=
distrans <= 2822:#md14
                            lu = 3
                        else:
                            if prf == 3 and 23 <= elv <= 166 and 12 <= slp <= 64 and 10 <= diswtr <= 20 and 2978
<= distrans <= 3252: #md15
                                lu = 5
                            else:
                                if prf == 3 and 210 <= elv <= 242 and 27 <= slp <= 41 and 198 <= diswtr <= 290
and 954 <= distrans <= 1038: #md16
                                    lu = 6
                    else:
                        if lu == 5 and prf == 3 and 105 <= elv <= 300 and 6 <= slp <= 73 and 0 <= diswtr <= 681
and 15 <= distrans <= 3118:#md17
                            lu = 102 #proacacia
                        else:
                            if lu == 6:
                                if prf == 1 and 20 <= elv <= 138 and 0 <= slp <= 59 and 36 <= diswtr <= 282 and
25 <= distrans <= 501: #md18
                                    lu = 102 #procacia
                                else:
                                    if prf == 1 and 75 <= elv <= 321 and 1 <= slp <= 75 and 20 <= diswtr <= 469
and 5 <= distrans <= 979: #md19
                                        lu = 102 #proacacia
                                    else:
                                        if prf == 3 and 117 <= elv <= 280 and 1 <= slp <= 66 and 0 <= diswtr <=
359 and 54 <= distrans <= 3220: #md20
                                             lu = 112 #to cassava
        else:
            if con1 == 0 and con2 == 0 and con3 == 1:      #The propfes may not need to compare the probability
of change in order to upgrade forest covers, however, it still needs to check prb for other conversion
                if operator.mod(time,5) == 0:              #Forest upgrade Logic: bamboo>poor>mix>medium>rich
and young>poor>mix>medium>rich
                    if lu == 2 and prf in (1,2,3,4) and 0 <= elv <= 400 and 0 <= slp <= 80 and 0 <= diswtr <=
1000 and 0 <= distrans <= 4000:
                        lu = 1 #Medium to rich
                    else:
                        if lu == 3 and prf in (1,2,3,4) and 0 <= elv <= 400 and 0 <= slp <= 80 and 0 <= diswtr
<= 1000 and 0 <= distrans <= 4000:
                            lu = 5 #poor to mix
                        else:
```

```
                            if lu == 5 and prf in(1,2,3,4) and 0 <= elv <= 400 and 0 <= slp <= 80 and 0 <=
diswtr <= 1000 and 0 <= distrans <= 4000:
                                lu = 2 #mix to medium
                            else:
                                if lu in (4,6) and prf in (1,2,3,4) and 0 <= elv <= 400 and 0 <= slp <= 80 and 0
<= diswtr <= 1000 and 0 <= distrans <= 4000:
                                    lu = 3 #bamboo to poor
                                else:
                                    if lu == 18 and prf in (1,2,3,4) and 0 <= elv <= 400 and 0 <= slp <= 80 and
0 <= diswtr <= 1000 and 0 <= distrans <= 4000:
                                        lu = 3 #Young to poor
                                    else:
                                        if lu == 9 and prf == 3 and 133 <= elv <= 134 and 2 <= slp <= 3 and 45
<= diswtr <= 51 and 35 <= distrans <= 53: #md9
                                            lu = 112 # to cassava
                                        else:
                                            if lu == 101: #this is propfes so acacia could be kept
                                                if prf == 4 and 132 <= elv <= 144 and 1 <= slp <= 31 and 80 <=
diswtr <= 349 and 79 <= distrans <= 205: #md10
                                                    lu = 6
                                                else:
                                                    if prf == 4 and 132 <= elv <= 132 and 1 <= slp <= 2 and 60
<= diswtr <= 70 and 88 <= distrans <= 99: #md11
                                                        lu = 112 #to cassava
                                            else:
                                                if lu in (111,112) and prf == 4 and 132 <= elv <= 134 and 2 <=
slp <= 3 and 59 <= diswtr <= 91 and 34 <= distrans <= 81: #md12
                                                    lu = 6
                                                else:
                                                    if lu == 17 and prf == 4 and 31 <= elv <= 130 and 50 <= slp
<= 74 and 0 <= diswtr <= 14 and 2562 <= distrans <= 3234: #md13
                                                        lu = 6
            else:
                if con1==0 and con2==0 and con3==0 and prb >= iniprb:   #This is normal conversion function,
runs when no special promotion is called
                    if operator.mod(time,5) == 0 and time < 10:
                        if lu == 2:
                            if  prf in (3,4) and 170 <= elv <= 242 and 21 <= slp <= 51 and 0 <= diswtr <= 251
and 994 <= distrans <= 2843:  #md1
                                lu = 3
                            else:
```

```
                                if prf in (3,4) and 151 <= elv <= 217 and 6 <= slp <= 50 and 10 <= diswtr <= 293
and 948 <= distrans <= 2992: #md2
                                    lu = 5
                                else:
                                    if prf in (1,2,3,4) and 23 <= elv <= 44 and 35 <= slp <= 70 and 0 <= diswtr
<= 10 and 3153 <= distrans <= 3247:#md3
                                        lu = 17
                        else:
                            if lu == 3:
                                if prf == 2 and 28 <= elv <= 267 and 24 <= slp <= 73 and 10 <= diswtr <= 344 and
136 <= distrans <= 2004: #md4
                                    lu = 5
                                else:
                                    if prf == 2 and 148 <= elv <= 248 and 3 <= slp <= 41 and 64 <= diswtr <= 338
and 30 <= distrans <= 2191: #md5
                                        lu = 6
                                    else:
                                        if prf == 2 and 144 <= elv <= 313 and 9 <= slp <= 38 and 30 <= diswtr <=
334 and 418 <= distrans <= 3297:#md6
                                            lu = 18
                            else:
                                if lu == 5 and prf == 2 and 141 <= elv <= 261 and 10 <= slp <= 39 and 0 <=
diswtr <= 423 and 5 <= distrans <= 2973: #md7
                                    lu = 6
                                else:
                                    if lu == 6 and prf == 1 and 128 <= elv <= 191 and 2 <= slp <= 34 and 36 <=
diswtr <= 322 and 14 <= distrans <= 555: #md8
                                        lu = 101 #Default to cashew
                                    else:
                                        if lu == 9 and prf == 3 and 133 <= elv <= 134 and 2 <= slp <= 3 and 45
<= diswtr <= 51 and 35 <= distrans <= 53: #md9
                                            lu = 112 #Default: to cassava
                                        else:
                                            if lu in (101,102):
                                                if prf == 4 and 132 <= elv <= 144 and 1 <= slp <= 31 and 80 <=
diswtr <= 349 and 79 <= distrans <= 205: #md10
                                                    lu = 6
                                                else:
                                                    if prf == 4 and 132 <= elv <= 132 and 1 <= slp <= 2 and 60
<= diswtr <= 70 and 88 <= distrans <= 99: #md11  Note: LU code 11 is divided into 111-Maize and 112-Cassava.
Conversion to annual crops will set default to 112-Cassava
```

```
                                                        lu = 112
                                            else:
                                                if lu in (111,112) and prf == 4 and 132 <= elv <= 134 and 2 <=
slp <= 3 and 59 <= diswtr <= 91 and 34 <= distrans <= 81: #md12
                                                    lu = 6
                                                else:
                                                    if lu == 17 and prf == 4 and 31 <= elv <= 130 and 50 <= slp
<= 74 and 0 <= diswtr <= 14 and 2562 <= distrans <= 3234: #md13
                                                        lu = 6
                    else:
                        if operator.mod(time,10) == 0:
                            if lu == 2:
                                if prf == 3 and 150 <= elv <= 235 and 15 <= slp <= 49 and 10 <= diswtr <= 36 and
2553 <= distrans <= 2822:#md14
                                    lu = 3
                                else:
                                    if prf == 3 and 23 <= elv <= 166 and 12 <= slp <= 64 and 10 <= diswtr <= 20
and 2978 <= distrans <= 3252: #md15
                                        lu = 5
                                    else:
                                        if prf == 3 and 210 <= elv <= 242 and 27 <= slp <= 41 and 198 <= diswtr
<= 290 and 954 <= distrans <= 1038: #md16
                                            lu = 6
                            else:
                                if lu == 5 and prf == 3 and 105 <= elv <= 300 and 6 <= slp <= 73 and 0 <= diswtr
<= 681 and 15 <= distrans <= 3118:#md17
                                    lu = 6
                                else:
                                    if lu == 6:
                                        if prf == 1 and 20 <= elv <= 138 and 0 <= slp <= 59 and 36 <= diswtr <=
282 and 25 <= distrans <= 501: #md18
                                            lu = 9
                                        else:
                                            if prf == 1 and 75 <= elv <= 321 and 1 <= slp <= 75 and 20 <= diswtr
<= 469 and 5 <= distrans <= 979: #md19
                                                lu = 101 #default: to cashew
                                            else:
                                                if prf == 3 and 117 <= elv <= 280 and 1 <= slp <= 66 and 0 <=
diswtr <= 359 and 54 <= distrans <= 3220: #md20
                                                     lu = 112   #Default to cassava
    return lu
```

```
#--------------------------------------------------------------------------------------------------

def profilecheck(cash,lab):
    if cash > 0 and lab > 0:              #Profile 1: call LULC change module: 0,3,8,18,19
        return 1
    if cash > 0 and lab <= 0:             #Profile 2: call LULC change module: 0,3,4,5,6,7
        return 2
    if cash <= 0 and lab > 0:             #profile 3: call LULC change module: 0,1,2,3,9,14,15,16,17,20
        return 3
    if cash < 0 and lab < 0:              #Profile 4: call LULC change module: 0,1,2,3,10,11,12,13
        return 4


#--------------------------------------------------------------------------------------------------
# Basic parameters
T = 10                                        #Cycles of simulation, can be input through
arcpy.GetParameterAsText(), Hugh and Crile recommend the cycle is 10 years because socioeconomic data and land
seems to be within 10 years
indir, inshp = ntpath.split(inpath)          #Set 2 new vars: indir and infile - result of the split
StartYear = inshp[7:11]                      #Extract starting year from input file, later using int() to convert
str to integer
basename = inshp[:-4]                        #[-4:] returns 'landuse2010'  [-8:] returns 'landuse', text to
compare with file name


#--------------------------------------
#PART III: Main simulation process

print 'STARTING SIMULATION...........'
for t in range (1, T + 1, 1):                           # SUPER LOOP:   The cycle of the iteration (10 years for
ex)
    print "Iteration no. :" + str(t)




#---------------------------------------
    for i in range (1, 182, 1):                      #Range 1-181 to cover all HH, this is the highest level of
loops, will iterate through each HH, used to update variables at HH level then use to iterate simulation
```

```
        expr = '"ID" = ' + str(i)                    #Only need "field" = 1 to pass to the Update cursor, for
every single HH


        #Calculate basic HH parameters at HH level:

        if t <> 1:
            print 'This is the ' + str(t) + ' year of HH no ' + str(i)
            rows = arcpy.UpdateCursor(inshp, expr)      #Use Update Cursor to calculate
            for row in rows:                            #Run through the selection of record one by one, if it
is not the first cycle, then vars need to be updated
                #Reset demographic vars back to 0: "What happened in last year ?"
                #When an iteration starts for each agent, all values extracted to LUS must be updated because
LUS could be changed from the previous cycle
                row.cashbal = 0     #To avoid huge accumulation at final year
                row.labbal = 0

                pfes, templab, templabneed, tempinv, tempprf, tempcash, jan, feb, mar, apr, may, jun, jul, aug,
sep, octorber, nov, dec = 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0

                bjan, bfeb, bmar, bapr, bmay, bjun, bjul, baug, bsep, boctorber, bnov, bdec = 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0


                if lowgrowth == 0:                      #Test the policy intervention: birth control or not
                    row.mouthfeed = row.mouthfeed*(1 + 0.021)            #Temp vars will hold the increased
number of population in each HH, this is simulation so vars can be a real with decimal
                    row.hhlab = row.hhlab*(1 + 0.0021)                        #Assume that labour force
increase at 10% of population growth rate 0.1*0.021 = 0.0021
                    row.agrlab = row.agrlab*(1 + 0.0021)
                else:
                    row.mouthfeed = row.mouthfeed*(1 + 0.015)            #Growth rate of population can be
soft-coded later
                    row.hhlab = row.hhlab*(1 + 0.0015)                       #Assume that labour force increase
at 10% of pop increase
                    row.agrlab = row.agrlab*(1 + 0.0015)

                row.nonagrlab = row.hhlab - row.agrlab       #Calculate the non-agriculture labour

                if highincome == 0:                     #No high income policy, keep increase at 5% as normal
pace
```

```
                row.nonagrinc = row.nonagrinc*(1 + 0.05)
            else:                                   #Highincome strategy: increase 10%
                row.nonagrinc = row.nonagrinc*(1 + 0.1)

            if propfes == 0:                        #No Pro PFES policy, payment keeps at 20/ha or 25/ha or
more if policy is promoted
                pfes = 20
            else:
                pfes = 25

            row.nonagrexp = row.nonagrexp*(1 + 0.05)   #Non agriculture expenditure of HH, 472NZD/head/year
(700k vnd/head/mth) at 5%

            #Update scenario values to the database: it will need for the conversion module because it use
those data as parameters
            row.procashew = procashew
            row.proacacia = proacacia
            row.propfes = propfes

            rows.updateRow(row)
        del row, rows



        #Income from agricultural activities: need to calculate by cell first then aggregate later to HH
level (=area of cell*parameters of each landuse type)
        rows = arcpy.UpdateCursor(inshp, expr)
        for row in rows:
            if row.CLU == 9:                            #Labour requirement and balance by HH then by each
cell
                row.CJAN = row.CFEB =  row.CJUN = row.CJUL = row.CAUG =  row.CNOV = row.CDEC =
row.CAREA*150/7*1 #Rice require 150 lab/ha/year,7 months, each month: 150/7. A agrlab can work 20lab/month,
total per month - required per month = balance per month
                row.CMAR = row.CAPR = row.CMAY = row.CSEP = row.COCT = row.CAREA*150/7*0    # *0 because
those months do not need labours
                row.CPRF = row.CAREA*(3529.41 - 1060.40)*0.6*0.7 #Income per ha: 3529.41 - Investment:
1060.41 = profit from agriculture. In general, HH can cultivate about 60% of their measured area and get 70%
from standard profit for many reasons
                row.CINV = row.CAREA*1060.40

            if row.CLU == 111:                          #Corn-Maize recoded as 111
```

```
                    row.CJAN = row.CFEB =  row.CJUN = row.CJUL = row.CAUG =  row.CNOV = row.CDEC = 
row.CAREA*100/7*1
                    row.CMAR = row.CAPR = row.CMAY = row.CSEP = row.COCT = row.CAREA*100/7*0
                    row.CPRF = row.CAREA*(2352.94 - 883.65)*0.6*0.7
                    row.CINV = row.CAREA*883.65

                if row.CLU == 112:                          #Cassava recoded as 112
                    row.CJAN = row.CFEB =  row.CJUN = row.CJUL = row.CAUG =  row.CNOV = row.CDEC = 
row.CAREA*130/7*1
                    row.CMAR = row.CAPR = row.CMAY = row.CSEP = row.COCT = row.CAREA*130/7*0
                    row.CPRF = row.CAREA*(3529.41 - 1147.71)*0.6*0.7
                    row.CINV = row.CAREA*1147.71

                if row.CLU == 101:                          #Cashew recoded as 101
                    row.CJAN = row.CFEB = row.CMAR = row.CAPR = row.CMAY = row.CJUN = row.CJUL = row.CAUG = 
row.CSEP = row.COCT = row.CNOV = row.CDEC = row.CAREA*50/12*1
                    row.CPRF = row.CAREA*(1035.29 - 441.41)*0.6*0.7
                    row.CINV = row.CAREA*441.41

                if row.CLU == 102 or row.CLU == 8:          #Acacia hybrid recoded as 102, plantation forest 
coded as 8
                    row.CJAN = row.CFEB = row.CMAR = row.CAPR = row.CMAY = row.CJUN = row.CJUL = row.CAUG = 
row.CSEP = row.COCT = row.CNOV = row.CDEC = row.CAREA*60/12*1
                    row.CPRF = row.CAREA*(1470.59 - 529.76)*0.6*0.7
                    row.CINV = row.CAREA*529.76

                if row.CLU in (1,2,3,4,5,6,7,18):           #Forestry: 
rich/medium/poor/young/mixed/evergreen/bamboo: $20 pay per ha and $100 of second forest product
                    row.CJAN = row.CFEB = row.CMAR = row.CAPR = row.CMAY = row.CJUN = row.CJUL = row.CAUG = 
row.CSEP = row.COCT = row.CNOV = row.CDEC = row.CAREA*10/12*1
                    row.CPRF = row.CAREA*((pfes + 100) - 88.24)
                    row.CINV = row.CAREA*88.24

                if row.CLU == 17 and row.aqua == 1:         #Aquaculture: land must be 17 and aqua = 1, 
calculate per 50x50m square of water surface
                    row.CJAN = row.CFEB = row.CMAR = row.CAPR = row.CMAY = row.CJUN = row.CJUL = row.CAUG = 
row.CSEP = row.COCT = row.CNOV = row.CDEC = row.CAREA*52/12*1/0.25
                    row.CPRF = row.CAREA*((1176.47 - 459.18))/0.25*0.6*0.7
                    row.CINV = row.CAREA*459.18/0.25

                rows.updateRow(row)
```

```
            del row, rows

            #Calculate the labour balance by month and by year in each HH:
            rows = arcpy.SearchCursor(inshp, expr)       #Only the SearchCursor update the household vars
exactly, the UpdateCursor does not accumulate values in a temp var like SearchCursor
            for row in rows:
                jan += row.CJAN
                feb += row.CFEB
                mar += row.CMAR
                apr += row.CAPR
                may += row.CMAY
                jun += row.CJUN
                jul += row.CJUL
                aug += row.CAUG
                sep += row.CSEP
                octorber += row.COCT
                nov += row.CNOV
                dec += row.CDEC

                bjan, bfeb, bmar, bapr, bmay, bjun, bjul, baug, bsep, boctorber, bnov, bdec = row.agrlab*20 -
jan, row.agrlab*20 - feb, row.agrlab*20 - mar, row.agrlab*20 - apr, row.agrlab*20 - may, row.agrlab*20 - jun,
row.agrlab*20 - jul, row.agrlab*20 - aug, row.agrlab*20 - sep, row.agrlab*20 - octorber, row.agrlab*20 - nov,
row.agrlab*20 - dec
                a = [bjan, bfeb, bmar, bapr, bmay, bjun, bjul, baug, bsep, boctorber, bnov, bdec]

                tempinv += row.CINV
                tempprf += row.CPRF

            del row, rows

            rows = arcpy.UpdateCursor(inshp, expr)        #Update value
            for row in rows:

                row.labbal = sum(a)                       #Labour balance of each HH through a temp var 'a'

                row.agrinv = tempinv

                row.agrprf = tempprf

                #Labour surplus from the agr labour (seasonal) could work to earn extra income
                if row.labbal > 0:
```

```
                    if highincome == 1:
                        row.extra = row.labbal*9*0.3        #Assume that they can work 30% of their seasonal
time: highincome $9 per day, low income $7
                    else:
                        row.extra = row.labbal*7*0.3         #Low income
                row.cashbal = row.agrprf +row.extra + row.nonagrinc - row.nonagrexp  #cashbal and labbal will be
used to calculate the profile of each HH: labbal is aggregated var then extra do not need to accumulate
                rows.updateRow(row)
            del row, rows


            rows = arcpy.UpdateCursor(inshp, expr)        #Update value
            for row in rows:
                #Calculate the probability of each cell following the econometric function, calculate the cdf
                row.CFUNC = -4.443067 - 1.463567*row.lastchange + 0.0119636*row.elv + 0.8775443*row.agrlab
#This equation took from chapter 4, lastchange in this case means change from 00-05, at the end of each
simulation, the change of each cell will be updated to serve as the lastchange of next simulation
                row.CPRB = norm.cdf(row.CFUNC)             #Calculate the probability of each change

                #Store the current land use value in each cell to quantify the change after each cycle
                row.CLUC = row.CLU

                rows.updateRow(row)
            del row, rows

            print 'HH demographical vars are updated'

    #----------------------------------------------------------------------

            #Third cursor (Update) will compare cashbal and labbal to classify HH to decision profiles:
            rows = arcpy.UpdateCursor(inshp, expr)  #'"ID" = 1' parsing the expr
            for row in rows:
                #Check the credit policy before calculate:
                if credit == 0:                           #No cash injection
                    row.cashrate = row.cashbal/(row.nonagrexp + row.agrinv)            #Calculate at the first
row of cursor is enough because all rows of a HH will be the same
                    row.labrate = row.labbal/(row.hhlab*20*12)          #Calculate at the first row of cursor is
enough because all rows of a HH will be the same

                else:                                        #Activate the credit policy
```

```
                    borrow = 300                            #HH with cashbal < -300 can borrow 300 for example ! 
Can be soft-coded to reflect the "interest rate" influence
                    if row.cashbal < (0- borrow):           #only HH with - cashbalance and < -300 cound borrow 
300
                        row.cashrate = (row.cashbal + borrow)/(row.nonagrexp + row.agrinv)            #Calculate 
at the first row of cursor is enough because all rows of a HH will be the same
                        row.labrate = row.labbal/(row.hhlab*20*12)          #Calculate at the first row of 
cursor is enough because all rows of a HH will be the same
                    else:
                        row.cashrate = row.cashbal/(row.nonagrexp + row.agrinv)            #Calculate at the 
first row of cursor is enough because all rows of a HH will be the same
                        row.labrate = row.labbal/(row.hhlab*20*12)

                row.profile = profilecheck(row.cashrate,row.labrate)              #A module that calculate the 
profile of HH based on cashrate, labrate

                row.CLU = 
conversion(row.procashew,row.proacacia,row.propfes,row.profile,row.CLU,row.elv,row.slp,row.diswtr,row.distrans,t
,row.CPRB,row.CINIPRB)

                if row.CLUC == row.CLU:
                    row.lastchange = 0
                else:
                    row.lastchange = 1

                rows.updateRow(row)

            del row, rows

        else:                                                       #Else means t = 1 the first year, should run 
the same code block for calculation and cell update
            print 'This is first year of HH no ' + str(i)

print 'SIMULATION ENDED!!!!'


#Stop timing
stoptime = time.time()
print 'Processing time: ' + str((stoptime - starttime)/60) + 'mins'
```